Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 18–39

SIGNED COMPOUND POISSON INTEGER-VALUED GARCH PROCESSES

E. GONÇALVES AND N. MENDES-LOPES

ABSTRACT: We propose signed compound Poisson integer-valued GARCH processes for the modelling of the difference of count time series data. We investigate the theoretical properties of these processes and we state their ergodicity and stationarity under mild conditions. We discuss the conditional maximum likelihood estimator when the series appearing in the difference are INGARCH with geometric distribution and explore its finite sample properties in a simulation study. Two real data examples illustrate this methodology.

KEYWORDS: Integer-valued time series, GARCH model, compound Poisson distributions.
AMS SUBJECT CLASSIFICATION (2010): 62M10.

1. Introduction

The practical relevance of count time series has led to the development of several class of integer-valued models in order to better describe and capture the main characteristics of this kind of data. Among these classes, we highlight that of the compound Poisson INGARCH (CP-INGARCH) models, introduced in Gonçalves et al (2015), that enlarges and unifies the main INGARCH processes present in literature and has the ability of capturing simultaneously characteristics of overdispersion and conditional heteroscedasticity, in a general distributional context.

Some papers emerged in the literature in recent years (a.e. Karlis and Ntzoufras, 2006, 2009; Koopman et al, 2014) have shown the interest of signed integer-valued time series defined by the Skellam distribution, which is constructed as differences in pairs of Poisson counts independent or not. Such models allow us to describe, for example, the differences over time in the number of accidents, catastrophes, or people contracting certain epidemic disease in two cities, two world regions, or two populations.

Several studies have shown that Poissonian models have some limitations to describe counting data, particularly because this kind of data usually has

Received October 1, 2018.

overdispersion characteristics, which are not captured by Poissonian distributions but by others such as negative binomial (NB), Neyman-type A, generalized Poisson (see a.e. Weiß, 2009; Zhu, 2011, 2012; Gonçalves et al, 2015) all of them included in the supra referred CP-INGARCH models. Therefore, to consider the differences of general counting models have obviously theoretical and practical interest.

Following this idea we present here a $\mathbb{Z}$-valued counting model defined as the difference between two general independent CP-INGARCH processes, which will allow to describe in practice such kind of data, even when the phenomena under study has different distributional behavior in each situation considered (in that difference). We observe that our proposal differs from those in which models to adjust time series with $\mathbb{Z}$-values are based on thinning operators, a subject where Kim and Park (2008) was pioneering; a good review of these thinning operators models is presented in Scotto et al (2015).

For sake of technical simplicity, we start with the study of a bivariate model defined by two independent CP-INGARCH processes from which the difference model, that is the signed CP-INGARCH one, is constructed. The main probabilistic properties of weak and strict stationarity and ergodicity are deduced for a wide class of these kind of bivariate models. After we consider any measurable function of the marginal processes and in what concerns the statistical analysis we concentrate our study in the estimation by the conditional maximum likelihood method of a particular case of signed geometric INGARCH model, that is, that corresponding to a bivariate model whose marginal processes are particular NB-INGARCH ones.

The remainder of the paper is organized as follows. In Section 2 we recall the class of CP-INGARCH models by presenting its definition and construction as well as its main subclasses. The bivariate process is introduced in Section 3. Properties of stationarity and ergodicity are developed and, in particular, necessary and sufficient conditions of weak and strict stationarity are established. The signed CP-INGARCH model is defined in Section 4. For the particular case of the signed geometric INGARCH processes, we develop in this Section the conditional maximum likelihood estimator of the model parameter vector. The performance of the conditional likelihood estimator in finite samples is evaluated via simulation experiments. We observe that with our models, we can compare a temporal phenomenon in different periods (sufficiently distant to assume independence) but also compare it in different situations; we illustrate it by discussing the monthly counts of poliomyelitis cases recorded in the

United States in two periods as well as another application related to the number of Olympic medals won by Swiss and Dutch athletes over all time. Some concluding remarks end the paper.

2. Preliminairies

Let us recall the definition of Compound Poisson integer-valued GARCH model introduced in Gonçalves et al (2015) as well as some results that are relevant for the study presented in next Sections.

Let $X = (X_t, t \in \mathbb{Z})$ be a stochastic process with values in $\mathbb{N}_0$ and, for any $t \in \mathbb{Z}$, let $\underline{X}_{t-1}$ be the $\sigma-$field generated by $\{X_{t-j}, j \geq 1\}$.

Definition 2.1 (CP-INGARCH($\mathbf{p}, \mathbf{q}$) model)**.** *The process X is said to satisfy a Compound Poisson INteger-valued GARCH model with orders p and q (p, $q \in \mathbb{N}$) if, $\forall t \in \mathbb{Z}$, the characteristic function of $X_t|\underline{X}_{t-1}$ is given by*

$$\begin{cases} \Phi_{X_t|\underline{X}_{t-1}}(u) = \exp\left\{i\frac{\lambda_t}{\varphi_t'(0)}\left[\varphi_t(u) - 1\right]\right\}, \ u \in \mathbb{R} \\ E(X_t|\underline{X}_{t-1}) = \lambda_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j X_{t-j} + \sum_{k=1}^{q} \beta_k \lambda_{t-k} \end{cases} \tag{1}$$

for some constants $\alpha_0 > 0$, $\alpha_j \geq 0$ ($j = 1, ..., p$), $\beta_k \geq 0$ ($k = 1, ..., q$), and where $(\varphi_t, t \in \mathbb{Z})$ is a family of characteristic functions on $\mathbb{R}$, $\underline{X}_{t-1}$-measurable associated to a family of discrete laws with support $\mathbb{N}_0$ and finite mean. i denotes the imaginary unit.

As $\varphi_t, t \in \mathbb{Z}$, is the characteristic function of a discrete law with support $\mathbb{N}_0$ and finite mean, the derivative of φ_t at $u = 0$, $\varphi_t'(0)$, exists and is nonzero.

In the previous definition, if $\beta_k = 0$, $k = 1, ..., q$, the CP-INGARCH(p, q) model is simply denoted CP-INARCH(p).

Remark 2.1. (1) *As the conditional distribution of X_t is a discrete compound Poisson law with support $\mathbb{N}_0$ then, $\forall t \in \mathbb{Z}$ and conditionally to $\underline{X}_{t-1}$, X_t can be identified in distribution as*

$$X_t \stackrel{d}{=} \sum_{j=1}^{N_t} X_{t,j}, \tag{2}$$

where N_t follows a Poisson law with parameter $\lambda_t^* = i\ \lambda_t/\varphi_t'(0)$, and $X_{t,1}, \ldots, X_{t,N_t}$ are discrete independent random variables, with support contained in $\mathbb{N}_0$, independent of N_t and having characteristic function φ_t with first derivative at zero, that is, with finite mean. We note that *the characteristic function φ_t being $\underline{X}_{t-1}$-mensurable may be a random function. This means that φ_t may depend on the previous observations of the process.*

(2) *Let us consider the polynomials*
$A(L) = \alpha_1 L + ... + \alpha_p L^p$ *and* $B(L) = 1 - \beta_1 L - ... - \beta_q L^q$,
where L is the backshift operator. To ensure the existence of the inverse of $B(L)$ we suppose that the roots of $B(z) = 0$ lie outside the unit circle which, for non-negative β_j, is equivalent to $\sum_{j=1}^{q} \beta_j < 1$. Under this assumption, the conditional expectation of the model (1) may be rewritten in the form

$$B(L)\lambda_t = \alpha_0 + A(L)X_t \Leftrightarrow \lambda_t = \alpha_0 B^{-1}(1) + B^{-1}(L)A(L)X_t$$

that is, with $B^{-1}(L)A(L) = \sum_{j=1}^{\infty} \psi_j L^j$,

$$\lambda_t = \alpha_0 B^{-1}(1) + \sum_{j=1}^{\infty} \psi_j X_{t-j},$$

which expresses a CP-INARCH($+\infty$) representation of the model (1).

(3) *The process X satisfying the model (1) is first order stationary if and only if $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$. Under this condition, the processes (X_t) and (λ_t) are both first order stationary and*

$$E(X_t) = E(\lambda_t) = \frac{\alpha_0}{1 - \sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j}.$$

(4) *If φ_t is deterministic, $t \in \mathbb{Z}$, and independent of t, there is a strictly and weakly stationary and ergodic process that satisfies the model (1), if and only if $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$.*

(5) *The sub-class of CP-INGARCH models with φ_t deterministic and independent of t is still quite vaste including, among others. the INGARCH (Ferland et al, 2006), Negative-Binomial DINARCH (Xu et al, 2012), Generalized Poisson INGARCH (Zhu, 2012), Neyman type-A and GEOMP2 (Gonçalves et al, 2015) INGARCH models.*

3. Bivariate model

3.1. Definition.

Let $X = (X_t, t \in \mathbb{Z})$ be a bivariate stochastic process, $X_t = (X_{1,t}, X_{2.t})$, where $X_1 = (X_{1,t}, t \in \mathbb{Z})$ and $X_2 = (X_{2,t}, t \in \mathbb{Z})$ are univariate processes and, for any $t \in \mathbb{Z}$, let $\underline{X}_{t-1}$ be the $\sigma-$field generated by $\{X_{t-j}, j \geq 1\}$. If X_1 satisfy a CP-INGARCH(p,q) model and X_2 satisfy a CP-INGARCH$(\widetilde{p},\widetilde{q})$ model and if, for any $t\in\mathbb{Z}$, $X_{1,t}$ e $X_{2.t}$ are independent relatively to the law conditioned by $\underline{X}_{t-1}$, then the characteristic function of $X_t|\underline{X}_{t-1}$ is, for any $t\in\mathbb{Z}$, given by

$$\begin{cases} \Phi_{X_t|\underline{X}_{t-1}}(u,v)=\exp\left\{i\frac{M_t}{\varphi_t'(0)}\left[\varphi_t(u)-1\right]\right\}\exp\left\{i\frac{\widetilde{M}_t}{\widetilde{\varphi}_t'(0)}\left[\widetilde{\varphi}_t(v)-1\right]\right\},\ (u,v)\in\mathbb{R}^2 \\ E(X_{1,t}|\underline{X}_{1,t-1})=M_t=\alpha_0+\sum_{j=1}^{p}\alpha_jX_{1,t-j}+\sum_{k=1}^{q}\beta_kM_{t-k}, \\ E(X_{2,t}|\underline{X}_{2,t-1})=\widetilde{M}_t=\widetilde{\alpha}_0+\sum_{j=1}^{\widetilde{p}}\widetilde{\alpha}_jX_{2,t-j}+\sum_{k=1}^{\widetilde{q}}\widetilde{\beta}_k\widetilde{M}_{t-k}, \end{cases} \tag{3}$$

for some constants $\alpha_0>0$, $\alpha_j\geq 0$ $(j=1,...,p)$, $\beta_k\geq 0$ $(k=1,...,q)$, $\widetilde{\alpha}_0>0$, $\widetilde{\alpha}_j\geq 0$ $(j=1,...,\widetilde{p})$, $\widetilde{\beta}_k\geq 0$ $(k=1,...,\widetilde{q})$ and where $(\varphi_t,t\in\mathbb{Z})$ and $(\widetilde{\varphi}_t,t\in\mathbb{Z})$ are two families of characteristic functions on $\mathbb{R}$, $\underline{X}_{1,t-1}$ and $\underline{X}_{2,t-1}-$measurable, respectively, each one associated to a family of discrete laws with support $\mathbb{N}_0$ and finite mean.

We assume in what follows the hypothesis

$$H_1:\sum_{k=1}^{q}\beta_k<1\ \text{ and }\ \sum_{k=1}^{\widetilde{q}}\widetilde{\beta}_k<1. \tag{4}$$

Introducing the polynomials

$A(L)=\alpha_1L+...+\alpha_pL^p$ and $B(L)=1-\beta_1L-...-\beta_qL^q$,

$\widetilde{A}(L)=\widetilde{\alpha}_1L+...+\widetilde{\alpha}_{\widetilde{p}}L^{\widetilde{p}}$ and $\widetilde{B}(L)=1-\widetilde{\beta}_1L-...-\widetilde{\beta}_{\widetilde{q}}L^{\widetilde{q}}$,

we may ensure the existence of the following representations for M_t and $\widetilde{M}_t$

$$M_t=\alpha_0B^{-1}(1)+\sum_{j=1}^{\infty}\psi_jX_{1,t-j},\qquad \widetilde{M}_t=\widetilde{\alpha}_0\widetilde{B}^{-1}(1)+\sum_{j=1}^{\infty}\widetilde{\psi}_jX_{2,t-j},$$

where ψ_j (resp., $\widetilde{\psi}_j$) is the coefficient of z^j in the Maclaurin expansion of $A(z)/B(z)$ (resp., $\widetilde{A}(z)/\widetilde{B}(z)$), that is, $B^{-1}(L)A(L)=\sum_{j=1}^{\infty}\psi_jL^j$ (resp., $\widetilde{B}^{-1}(L)\widetilde{A}(L)=\sum_{j=1}^{\infty}\widetilde{\psi}_jL^j$).

3.2. First and second order stationarity.

The process X is first order stationary if and only if $\sum_{i=1}^{p}\alpha_i+\sum_{j=1}^{q}\beta_j<1$ and $\sum_{i=1}^{\widetilde{p}}\widetilde{\alpha}_i+\sum_{j=1}^{\widetilde{q}}\widetilde{\beta}_j<1$. Under these conditions, the processes $(X_{1,t})$ and

(M_t) are both first order stationary, as well as $(X_{2,t})$ and $(\widetilde{M}_t)$, and we have

$$\begin{aligned} E(X_{1,t}) &= E(M_t) = \mu = \frac{\alpha_0}{1-\sum_{i=1}^{p}\alpha_i - \sum_{j=1}^{q}\beta_j}, \\ E(X_{2,t}) &= E(\widetilde{M}_t) = \widetilde{\mu} = \frac{\widetilde{\alpha}_0}{1-\sum_{i=1}^{\widetilde{p}}\widetilde{\alpha}_i - \sum_{j=1}^{\widetilde{q}}\widetilde{\beta}_j}. \end{aligned}$$

The study of the second order stationarity of CP- INGARCH models was undertaken in Gonçalves et al (2015) under the condition

$$H_2 : -i\frac{\varphi_t^{''}(0)}{\varphi_t^{'}(0)} = v_0 + v_1\lambda_t,$$

with $v_0 \geq 0, v_1 \geq 0$, not simultaneously zero. The stated results are valuable for a quite general subclass including both random and deterministic characteristic functions φ_t; in particular, they apply to the INGARCH (Ferland et al, 2006), NB-INARCH (Zhu, 2011), NB-DINARCH (Xu et al, 2012), Generalized Poisson INGARCH (Zhu, 2012), Neyman type-A INGARCH models and also to all the models where the φ_t functions are determinist and independent of t. This study apply naturally to the bivariate model $X_t = (X_{1,t}, X_{2.t})$ since the matrices of variances-covariances, $\Gamma_{X_t}(h), h = 0, 1, 2, ...$, are given by

$$\begin{aligned} \Gamma_{X_t}(h) &= \begin{bmatrix} Cov\left(X_{1,t+h}, X_{1,t}\right) & Cov\left(X_{1,t+h}, X_{2,t}\right) \\ Cov\left(X_{1,t+h}, X_{2,t}\right) & Cov\left(X_{2,t+h}, X_{2,t}\right) \end{bmatrix} \\ &= \begin{bmatrix} Cov\left(X_{1,t+h}, X_{1,t}\right) & 0 \\ 0 & Cov\left(X_{2,t+h}, X_{2,t}\right) \end{bmatrix} \\ &= \begin{bmatrix} \Gamma_{X_{1,t}}(h) & 0 \\ 0 & \Gamma_{X_{2,t}}(h) \end{bmatrix}, \end{aligned}$$

due to the independence of X_1 and X_2.

From Theorems 2 of Gonçalves et al (2015) and 3.1 of Gonçalves et al (2016), a necessary and sufficient condition of weak stationarity of processes $(X_{1,t})$ and $(X_{2.t})$ and consequently of (X_t) is deduced.

We refer, for example, that for a first order stationary CP-INGARCH(1,1) model verifying H_2, a necessary and sufficient condition of weak stationarity is $(\alpha_1 + \beta_1)^2 + v_1\alpha_1^2 < 1$. We also note that the autocovariances of a second order CP-INGARCH(p, q) process X and those of λ (respectively, Γ and $\widetilde{\Gamma}$) verify the linear equations

$$\Gamma(h)=\sum_{j=1}^{p}\alpha_j\Gamma(h-j)+\sum_{k=1}^{\min(h-1,q)}\beta_k\Gamma(h-k)+\sum_{k=h}^{q}\beta_k\widetilde{\Gamma}(h-k),\quad h\geq 1,$$

$$\widetilde{\Gamma}(h)=\sum_{j=1}^{\min(h,p)}\alpha_j\widetilde{\Gamma}(h-j)+\sum_{j=h+1}^{p}\alpha_j\Gamma(j-h)+\sum_{k=1}^{q}\beta_k\widetilde{\Gamma}(h-k),\quad h\geq 0,$$

assuming that $\sum_{k=h}^{q}\beta_k\widetilde{\Gamma}(h-k)=0$ if $h>q$ and $\sum_{j=h+1}^{p}\alpha_j\Gamma(j-h)=0$ if $h>p$. In particular, the autocovariances of a weakly stationary CP-INGARCH(1,1) are given by

$$\Gamma(h)=\frac{\alpha_1(1-\beta_1(\alpha_1+\beta_1))(\alpha_1+\beta_1)^{h-1}}{1-(\alpha_1+\beta_1)^2+\alpha_1^2}\Gamma(0),\qquad h\geq 1,$$

with, under H_2,

$$\Gamma(0)=\mu\frac{(v_0+v_1\mu)\left[1-(\alpha_1+\beta_1)^2+\alpha_1^2\right]}{1-(\alpha_1+\beta_1)^2-v_1\alpha_1^2}$$

where $\mu=\frac{\alpha_0}{1-\alpha_1-\beta_1}$.

3.3. Strict stationarity.

In this section we study the existence of strictly stationary solutions for the class of models introduced in (3). Following Ferland et al (2006) and Gonçalves et al (2015) we begin by building a first order stationary process solution of the bivariate model that, under certain conditions, will be strictly stationary and ergodic.

3.3.1. *Construction of a process solution when φ_t and $\widetilde{\varphi}_t$ are deterministic.*

Let us consider model (3) associated to a given family of characteristic functions $(\varphi_t,\widetilde{\varphi}_t,t\in\mathbb{Z})$ such that the hypothesis H_1 is satisfied. We assume

$$H_3:\varphi_t\ \text{ and }\ \widetilde{\varphi}_t\ \text{ are deterministic.}\tag{5}$$

Let $(U_t,t\in\mathbb{Z})$ be a sequence of independent real random variables distributed according to a discrete compound Poisson law with characteristic function

$$\Phi_{U_t}(u)=\exp\left\{\frac{\alpha_0}{B(1)}\frac{i}{\varphi_t'(0)}\left[\varphi_t(u)-1\right]\right\}.$$

For each $t \in \mathbb{Z}$ and $k \in \mathbb{N}$, let $Z_{t,k} = \{Z_{t,k,j}\}_{j\in\mathbb{N}}$ be a sequence of independent discrete compound Poisson random variables with characteristic function

$$\Phi_{Z_{t,k,j}}(u) = \exp\left\{\psi_k \frac{i}{\varphi'_{t+k}(0)}\left[\varphi_{t+k}(u) - 1\right]\right\},$$

where $(\psi_j, j \in \mathbb{N})$ is the sequence of coefficients associated to the CP-INARCH($+\infty$) representation of the model $X_{1,t}$. We note that $E(U_t) = \alpha_0 B^{-1}(1) = \psi_0$, $E(Z_{t,k,j}) = \psi_k$ and that $Z_{t,k,j}$ are identically distributed for each $(t,k) \in \mathbb{Z}\times\mathbb{N}$. We also assume that all the variables U_s, $Z_{t,k,j}$, $s,t \in \mathbb{Z}, k,j \in \mathbb{N}$, are mutually independent. Based on these random variables, we define the sequence $X_{1,t}^{(n)}$ as follows:

$$X_{1,t}^{(n)} = \begin{cases} 0, & n < 0 \\ U_t, & n = 0 \\ U_t + \sum_{k=1}^{n}\sum_{j=1}^{X_{1,t-k}^{(n-k)}} Z_{t-k,k,j}, & n > 0 \end{cases}, \tag{6}$$

where it is assumed that $\sum_{j=1}^{0} Z_{t-k,k,j} = 0$.

We introduce, analogously, the sequences, independent of the previous ones, $(\widetilde{U}_t, t \in \mathbb{Z})$, $\widetilde{Z}_{t,k} = \{Z_{t,k,j}\}_{j\in\mathbb{N}}$ for each $t \in \mathbb{Z}$ and $k \in \mathbb{N}$, with characteristic functions

$$\begin{aligned} \Phi_{\widetilde{U}_t}(u) &= \exp\left\{\frac{\widetilde{\alpha}_0}{\widetilde{B}(1)}\frac{i}{\widetilde{\varphi}'_t(0)}\left[\widetilde{\varphi}_t(u) - 1\right]\right\}, \\ \Phi_{\widetilde{Z}_{t,k,j}}(u) &= \exp\left\{\widetilde{\psi}_k \frac{i}{\widetilde{\varphi}'_{t+k}(0)}\left[\widetilde{\varphi}_{t+k}(u) - 1\right]\right\} \end{aligned}$$

and define the sequence $X_{2,t}^{(n)}$ as:

$$X_{2,t}^{(n)} = \begin{cases} 0, & n < 0 \\ \widetilde{U}_t, & n = 0 \\ \widetilde{U}_t + \sum_{k=1}^{n}\sum_{j=1}^{X_{2,t-k}^{(n-k)}} \widetilde{Z}_{t-k,k,j}, & n > 0 \end{cases}. \tag{7}$$

In what follows we present some properties of the sequence

$$\left(X_{1,t}^{(n)}, X_{2,t}^{(n)}, n \in \mathbb{N}\right),$$

that are direct consequences of Ferland et al (2006) and Gonçalves et al (2015).

Property 3.1. *If* $\sum_{i=1}^{p}\alpha_i+\sum_{j=1}^{q}\beta_j<1$ *and* $\sum_{i=1}^{\widetilde{p}}\widetilde{\alpha}_i+\sum_{j=1}^{\widetilde{q}}\widetilde{\beta}_j<1$ *then* $\{(X_{1,t}^{(n)},X_{2,t}^{(n)},t\in\mathbb{Z}),\ n\in\mathbb{Z}\}$ *is a sequence of first order stationary processes such that, as* $n\to\infty$,

$$\left(E\left(X_{1,t}^{(n)}\right),E\left(X_{2,t}^{(n)}\right)\right)\longrightarrow(\mu,\widetilde{\mu}).$$

Property 3.2. *If* $\sum_{i=1}^{p}\alpha_i+\sum_{j=1}^{q}\beta_j<1$, $\sum_{i=1}^{\widetilde{p}}\widetilde{\alpha}_i+\sum_{j=1}^{\widetilde{q}}\widetilde{\beta}_j<1$ *and* φ_t *and* $\widetilde{\varphi}_t$ *are derivable at zero up to order 2, then the sequence* $\{(X_{1,t}^{(n)},X_{2,t}^{(n)},t\in\mathbb{Z}),\ n\in\mathbb{Z}\}$ *converges almost surely, in* L^1 *and* L^2 *to a process* $(X_1^*,X_2^*)=(X_{1,t}^*,X_{2,t}^*,t\in\mathbb{Z})$.

Taking into account the previous results, we obtain the next lemma that will be useful to establish the existence of a strictly stationary and ergodic process satisfying the bivariate model (3).

Lemma 3.1. *Under the hypothesis* H_3*, the process* $(X_1^*,\ X_2^*)$ *is a solution of the model if* $\sum_{j=1}^{p}\alpha_j+\sum_{k=1}^{q}\beta_k<1$ *and* $\sum_{i=1}^{\widetilde{p}}\widetilde{\alpha}_i+\sum_{j=1}^{\widetilde{q}}\widetilde{\beta}_j<1$.

***Proof*.** The almost sure limit of the sequence $(X_{1,t}^{(n)},X_{2,t}^{(n)})$ is a solution of the model (3) since, for $(u,v)\in\mathbb{R}^2$,

$$\Phi_{\left(X_{1,t}^*|\underline{X}_{1,t-1}^*,X_{2,t}^*|\underline{X}_{2,t-1}^*\right)}(u,v)=\lim_{n\to+\infty}\Phi_n(u,v)\tag{8}$$

$$=\exp[i\frac{M_t^*}{\varphi'(0)}(\varphi(u)-1)]\exp[i\frac{\widetilde{M_t^*}}{\widetilde{\varphi}'(0)}(\widetilde{\varphi}(v)-1)],\tag{9}$$

with Φ_n the characteristic function of the sequence $\left(r_t^{(n)}|\underline{X}_{1,t-1}^*,\widetilde{r}_t^{(n)}|\underline{X}_{2,t-1}^*\right)$, where

$$r_t^{(n)}=U_t+\sum_{k=1}^{n}\sum_{j=1}^{X_{1,t-k}^*}Z_{t-k,k,j},\qquad \widetilde{r}_t^{(n)}=\widetilde{U}_t+\sum_{k=1}^{n}\sum_{j=1}^{X_{2,t-k}^*}\widetilde{Z}_{t-k,k,j}\tag{10}$$

and $M_t^*=\alpha_0B^{-1}(1)+\sum_{j=1}^{\infty}\psi_jX_{1,t-j}^*$, $\qquad\widetilde{M_t^*}=\widetilde{\alpha}_0\widetilde{B}^{-1}(1)+\sum_{j=1}^{\infty}\widetilde{\psi}_jX_{2,t-j}^*$.

As in Ferland et al (2006) and Gonçalves et al (2015), the equality in (8) follows from Paul Lévy theorem since, for a fixed t, the sequence $Y_{1,t}^{(n)}=r_t^{(n)}-X_{1,t}^{(n)}$ converges in mean to zero, when $n\to\infty$. So $Y_{1,t}^{(n)}$ and $X_{1,t}^*-X_{1,t}^{(n)}$ converge in probability to zero and

$$X_{1,t}^*-r_t^{(n)}=(X_{1,t}^*-X_{1,t}^{(n)})+(X_{1,t}^{(n)}-r_t^{(n)})=(X_{1,t}^*-X_{1,t}^{(n)})-Y_{1,t}^{(n)},$$

which allows to conclude that the sequence $r_t^{(n)}$ converges in probability to $X^*_{1,t}$ and then $r_t^{(n)}|\underline{X}^*_{1,t-1}$ converges in law to $X^*_{1,t}|\underline{X}^*_{1,t-1}$. In an analogous way, we conclude that $\widetilde{r}_t^{(n)}|\underline{X}^*_{2,t-1}$ converges in law to $X^*_{2,t}|\underline{X}^*_{2,t-1}$.

Let us obtain Φ_n. Conditionally to $\underline{X}^*_{1,t-1}$, we have

$$\Phi_{\sum_{j=1}^{X^*_{1,t-k}} Z_{t-k,k,j}}(u) = \prod_{j=1}^{X^*_{1,t-k}} \Phi_{Z_{t-k,k,j}}(u) = \exp\left\{ \sum_{j=1}^{X^*_{1,t-k}} \psi_k \frac{i}{\varphi_t'(0)} \left[\varphi_t(u) - 1\right] \right\}$$

$$= \exp\left\{ \psi_k X^*_{1,t-k} \frac{i}{\varphi_t'(0)} \left[\varphi_t(u) - 1\right] \right\},$$

and conditionally to $\underline{X}^*_{2,t-1}$, we have

$$\Phi_{\sum_{j=1}^{X^*_{2,t-k}} \widetilde{Z}_{t-k,k,j}}(v) = \prod_{j=1}^{X^*_{2,t-k}} \Phi_{\widetilde{Z}_{t-k,k,j}}(v) = \exp\left\{ \widetilde{\psi}_k X^*_{2,t-k} \frac{i}{\widetilde{\varphi}_t'(0)} \left[\widetilde{\varphi}_t(v) - 1\right] \right\}.$$

From the independence of the variables involved in the definition of $r_t^{(n)}$ and $\widetilde{r}_t^{(n)}$, we obtain

$$\begin{aligned}
\Phi_n(u,v) &= \exp\left(\frac{\alpha_0}{B(1)} \frac{i}{\varphi_t'(0)} \left[\varphi_t(u) - 1\right] + \sum_{k=1}^{n} \psi_k X^*_{1,t-k} \frac{i}{\varphi_t'(0)} \left[\varphi_t(u) - 1\right] \right) \times \\
&\quad \times \exp\left(\frac{\widetilde{\alpha}_0}{\widetilde{B}(1)} \frac{i}{\widetilde{\varphi}_t'(0)} \left[\widetilde{\varphi}_t(v) - 1\right] + \sum_{k=1}^{n} \widetilde{\psi}_k X^*_{2,t-k} \frac{i}{\widetilde{\varphi}_t'(0)} \left[\widetilde{\varphi}_t(v) - 1\right] \right) \\
&= \exp\left\{ \left(\frac{\alpha_0}{B(1)} + \sum_{k=1}^{n} \psi_k X^*_{1,t-k} \right) \frac{i}{\varphi_t'(0)} \left[\varphi_t(u) - 1\right] \right\} \times \\
&\quad \times \exp\left\{ \left(\frac{\widetilde{\alpha}_0}{\widetilde{B}(1)} + \sum_{k=1}^{n} \widetilde{\psi}_k X^*_{2,t-k} \right) \frac{i}{\widetilde{\varphi}_t'(0)} \left[\widetilde{\varphi}_t(v) - 1\right] \right\},
\end{aligned}$$

and thus, when $n \to \infty$, we have the equality presented in (9). ■

Remark 3.1. *As a consequence of Property 3.1 and previous lemma, the process (X_1^*, X_2^*) is, under the hypothesis H_3, a first order stationary solution of the model if $\sum_{j=1}^{p} \alpha_j + \sum_{k=1}^{q} \beta_k < 1$ and $\sum_{i=1}^{\widetilde{p}} \widetilde{\alpha}_i + \sum_{j=1}^{\widetilde{q}} \widetilde{\beta}_j < 1$.*

Now, we consider, additionally to the hypothesis H_3, that φ_t and $\widetilde{\varphi}_t$ are independent of t. In this subclass, it is possible to establish the strict stationarity and ergodicity of (X_1^*, X_2^*).

Theorem 3.1. *Let us consider the bivariate model defined by* (3) *with* φ_t *and* $\widetilde{\varphi}_t$, $t \in \mathbb{Z}$, *deterministic and independent of* t.

- **(a):** $\{(X_{1,t}^{(n)}, X_{2,t}^{(n)}, t \in \mathbb{Z}), n \in \mathbb{Z}\}$ *is a sequence of strictly stationary and ergodic processes.*
- **(b):** *There is a strictly stationary and ergodic process in* L^1 *that satisfies the bivariate model if and only if* $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ *and* $\sum_{i=1}^{\widetilde{p}} \widetilde{\alpha}_i + \sum_{j=1}^{\widetilde{q}} \widetilde{\beta}_j < 1$. *Moreover, its first two moments are finite.*

Proof. (a) The sequence $\{(X_{1,t}^{(n)}, t \in \mathbb{Z}), n \in \mathbb{Z}\}$ is strictly stationary since the sequences $(U_t, t \in \mathbb{Z})$ and $(Z_{t,k}, t \in \mathbb{Z}, k \in \mathbb{N})$, defined in Section 3.3.1, are in this case ($\varphi_t = \varphi$) of i.i.d. random variables. Moreover, $(X_{1,t}^{(n)})$ is a sequence of ergodic processes, because it is a measurable function of the sequence of i.i.d. random variables $\{(U_t, Z_{t,j}), t \in \mathbb{Z}, j \in \mathbb{N}\}$ (Durrett, 2010). Analogously, we conclude that the sequence $\{(X_{2,t}^{(n)}, t \in \mathbb{Z}), n \in \mathbb{Z}\}$ is strictly stationary and ergodic as it envolves the sequences $(\widetilde{U}_t, t \in \mathbb{Z})$ and $(\widetilde{Z}_{t,k}, t \in \mathbb{Z}, k \in \mathbb{N})$.

As $X_{1,t}^{(n)}$ and $X_{2,t}^{(n)}$ are independent for each t, the process $(X_{1,t}^{(n)}, X_{2,t}^{(n)}, t \in \mathbb{Z})$ is strictly stationary.

(b) In Lemma 3.1 we proved that $(X_{1,t}^*, X_{2,t}^*, t \in \mathbb{Z})$ is a solution of (3). So, it is enough to prove that when φ_t and $\widetilde{\varphi}_t$ are deterministic and independent of t, the almost sure limit is strictly stationary and ergodic. From (a), $(X_{1,t}^{(n)}, X_{2,t}^{(n)}, n \in \mathbb{Z})$ is a sequence of strictly stationary processes. Otherwise, $(X_{1,t}^{(n)}, X_{2,t}^{(n)})$ converges almost surely to $(X_{1,t}^*, X_{2,t}^*)$ if $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ and $\sum_{i=1}^{\widetilde{p}} \widetilde{\alpha}_i + \sum_{j=1}^{\widetilde{q}} \widetilde{\beta}_j < 1$. So, considering without loss of generality, the indexes $\{1, ..., k\}$, we have when n tends to $+\infty$

$$\begin{aligned} ((X_{1,1}^{(n)}, X_{2,1}^{(n)}), ..., (X_{1,k}^{(n)}, X_{2,k}^{(n)})) &\rightarrow \left((X_{1,1}^*, X_{2,1}^*), ..., (X_{1,k}^*, X_{2,k}^*)\right), \\ ((X_{1,1+h}^{(n)}, X_{2,1+h}^{(n)}), ..., (X_{1,k+h}^{(n)}, X_{2,k+h}^{(n)})) &\rightarrow ((X_{1,1+h}^*, X_{2,1+h}^*), ..., (X_{1,k+h}^*, X_{2,k+h}^*)), \end{aligned}$$

almost surely, for any $h \in \mathbb{Z}$, and consequently, in law. Considering the strict stationarity of $(X_{1,t}^{(n)}, X_{2,t}^{(n)})$ and the limit unicity, we conclude that $(X_{1,t}^*, X_{2,t}^*)$ is a strictly stationary process. Moreover, taking into account that $(X_{1,t}^{(n)}, X_{2,t}^{(n)})$ is the measurable function of $(U_t, Z_{t,j}, \widetilde{U}_t, \widetilde{Z}_{t,j})$ referred above, $(X_{1,t}^*, X_{2,t}^*)$ is the almost sure limit of a sequence of measurable functions, so a measurable one (Halmos, 1974), of the ergodic process $(U_t, Z_{t,j}, \widetilde{U}_t, \widetilde{Z}_{t,j})$. Thus $(X^*_{1,t}, X^*_{2,t})$ is ergodic.

Regarding the necessary condition of existence of a strictly stationary solution, we observe that if $(X_{1,t}, X_{2,t})$ is such a solution of the bivariate model, it is also first order stationary as, by hypothesis, it is a process of L^1. So, by Section 2 we have $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ and $\sum_{i=1}^{\widetilde{p}} \widetilde{\alpha}_i + \sum_{j=1}^{\widetilde{q}} \widetilde{\beta}_j < 1$. ∎

Remark 3.2. *Under the conditions of the previous theorem it follows that* $\left(X^*_{1,t}, X^*_{2,t}, t \in \mathbb{Z}\right)$ *is also a weakly stationary solution of the model* (3) *because it is a strictly stationary second order process, from Property 3.2 .*

4. Signed compound Poisson INGARCH models

If we consider that the law de $X_{1,t}|\underline{X}_{1,t-1}$ is any discrete Compound Poisson law, and analogously for $X_{2,t}|\underline{X}_{2,t-1}$, we see that the class of models proposed in the previous Section includes inumerous bivariate cases. As examples of INGARCH models related to discrete compound Poisson laws recently studied, we refer the Binomial negative (Zhu, 2011), generalized Poisson (Zhu, 2012), dispersed INARCH (Xu et al, 2012), geometric and Neyman type-A (Gonçalves et al, 2015), among others.

The study of the resulting model for any measurable function of $X_{1,t}$ e $X_{2,t}$ will be more or less complexe according to the retained laws. In particular, there is a strictly stationary and ergodic solution for the model if the conditional laws relative to $X_{1,t}$ and $X_{2,t}$ are chosen among INGARCH, NB-DINARCH, GP, Neyman Type-A and GEOMP2 models as, in these cases, the characteristic function of the compounding variables is deterministic and independent of t.

A natural transformation is the process difference $D_t = X_{1,t} - X_{2,t}$, $t \in \mathbb{Z}$.

Example 4.1. *Let us consider that* $X_{1,t}|\underline{X}_{1,t-1}$ *is Poisson distributed with parameter* $M_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j X_{1,t-j} + \sum_{k=1}^{q} \beta_k M_{t-k}$. *This model was introduced by Ferland et al (2006), is denoted INGARCH model and it belongs to CP-INGARCH models as it is enough to consider, in Observation 2.1,* φ_t *equal to the characteristic function of Dirac law concentrated in* $\{1\}$ *and* N_t *Poisson distributed with parameter* M_t. *Analogously, let us consider that the law of* $X_{2,t}|\underline{X}_{2,t-1}$ *is Poisson with parameter* $\widetilde{M}_t = \widetilde{\alpha}_0 + \sum_{j=1}^{\widetilde{p}} \widetilde{\alpha}_j X_{2,t-j} + \sum_{k=1}^{\widetilde{q}} \widetilde{\beta}_k \widetilde{M}_{t-k}$. *The constants verify* $\alpha_0 > 0$, $\alpha_j \geq 0$ $(j = 1, ..., p)$, $\beta_k \geq 0$ $(k = 1, ..., q)$, $\widetilde{\alpha}_0 > 0$, $\widetilde{\alpha}_j \geq 0$ $(j = 1, ..., \widetilde{p})$, $\widetilde{\beta}_k \geq 0$ $(k = 1, ..., \widetilde{q})$ *and*

$\sum_{i=1}^{p}\alpha_i+\sum_{j=1}^{q}\beta_j<1$ *and* $\sum_{i=1}^{\widetilde{p}}\widetilde{\alpha}_i+\sum_{j=1}^{\widetilde{q}}\widetilde{\beta}_j<1$. *The conditional law of the difference process,* $D_t=X_{1,t}-X_{2,t}$, $t\in\mathbb{Z}$, *is a Skellam law (Skellam, 1946), that is, with probability function given by*

$$P\left(D_t=d\mid \underline{D}_{t-1}\right)=\exp\left[-\left(M_t+\widetilde{M_t}\right)\right]\left(\frac{M_t}{\widetilde{M_t}}\right)^{\frac{d}{2}}I_{|d|}\left(2\sqrt{M_t\widetilde{M_t}}\right),d\in\mathbb{Z},$$

where $I_{|d|}(.)$ *is the modified Bessel fonction of the first kind, that is, (Abramowitz and Stegun, 1965)*

$$\begin{aligned}I_{\alpha}(x)&=\sum_{m=0}^{+\infty}\frac{1}{m!\Gamma(m+\alpha+1)}\left(\frac{x}{2}\right)^{2m+\alpha}, \text{when } \alpha \text{ is not an integer}\\ I_k(x)&=\lim_{\alpha\to k}I_{\alpha}(x), \text{when } k \text{ is an integer.}\end{aligned}$$

The process difference $D_t=X_{1,t}-X_{2,t}$, $t\in\mathbb{Z}$, *is an INGARCH process with values in* $\mathbb{Z}$ *and the previous study shows that it is strictly stationary and ergodic. We note that the Skellam law has a recognized utility in the modelling of the number of accidents (or murders, strikes, catastrophes, ...) registered, for example, in two towns, two populations, two years, ...*

We present now the first and second-order moments of the difference process. The conditional mean of D_t is

$$E\left(D_t|\underline{D}_{t-1}\right)=M_t-\widetilde{M_t},$$

and so the first unconditional moment is

$$E(D_t)=E\left(M_t-\widetilde{M_t}\right)=\frac{\alpha_0}{1-\sum_{i=1}^{p}\alpha_i-\sum_{j=1}^{q}\beta_j}-\frac{\widetilde{\alpha}_0}{1-\sum_{i=1}^{\widetilde{p}}\widetilde{\alpha}_i-\sum_{j=1}^{\widetilde{q}}\widetilde{\beta}_j}.$$

Moreover, $\Gamma_{D_t}(h)=\Gamma_{X_{1,t}}(h)+\Gamma_{X_{2,t}}(h)$, $h\in\mathbb{Z}$, due to the independence.

In particular, if $M_t=\alpha_0+\alpha_1X_{1,t-1}$ and $\widetilde{M_t}=\widetilde{\alpha}_0+\widetilde{\alpha}_1X_{2,t-1}$ we have

$$\begin{aligned}E(D_t)&=\frac{\alpha_0}{1-\alpha_1}-\frac{\widetilde{\alpha}_0}{1-\widetilde{\alpha}_1},\\ \Gamma_{D_t}(h)&=\alpha_1^h\,\Gamma_{X_{1,t}}(0)+\widetilde{\alpha}_1^h\,\Gamma_{X_{2,t}}(0),\quad h\geq 1\\ \Gamma_{X_{1,t}}(0)&=\frac{\alpha_0}{1-\alpha_1}\frac{\nu_0+\upsilon_1\frac{\alpha_0}{1-\alpha_1}}{1-(1+\upsilon_1)\alpha_1^2}\end{aligned}$$

and analogously for $\Gamma_{X_{2,t}}(0)$, with α_0, α_1, ν_0 and υ_1 replaced by $\widetilde{\alpha}_0$, $\widetilde{\alpha}_1, \widetilde{\nu}_0$ and $\widetilde{\upsilon}_1$.

Let us illustrate the study of the process difference when $X_{1,t}|\underline{X}_{1,t-1}$ and $X_{2,t}|\underline{X}_{2,t-1}$ are geometrically distributed, observing that the characteristic functions of the compounding variables are not deterministic in this case.

4.1. The signed geometric INGARCH model.

4.1.1. *Preliminaires*.

The geometric law belongs to the Compound Poisson distributions. In what follows we consider that $X_{1,t}|\underline{X}_{1,t-1}$ e $X_{2,t}|\underline{X}_{2,t-1}$ are geometrically distributed and we analyse some properties of the process difference.

Let us recall that if the random variable X is geometrically distributed with parameter p_1, that is, with probability function $P(X=k) = p_1(1-p_1)^k$, $k \in \mathbb{N}_0$, then its characteristic function is given by $\varphi(t) = \frac{p_1}{1-(1-p_1)\exp(it)}$, $t \in \mathbb{R}$, and we have, for example, $E(X) = \frac{1-p_1}{p_1}$, $V(X) = \frac{1-p_1}{p_1^2}$. Moreover, if Y is another random variable that is independent of X and following a geometric law with parameter p_2, then the difference $X-Y$ has support $\mathbb{Z}$ and probability function

$$
\begin{aligned}
P(X-Y=k) &= \sum_{n=0}^{+\infty} P(X=n+k, Y=n) \\
&= \begin{cases} \dfrac{p_1 p_2}{1-(1-p_1)(1-p_2)}(1-p_1)^k, & k \in \mathbb{N}_0 \\ \dfrac{p_1 p_2}{1-(1-p_1)(1-p_2)}(1-p_2)^{-k}, & -k \in \mathbb{N}. \end{cases}
\end{aligned}
$$

The geometric INGARCH process $(X_t, t \in \mathbb{Z})$ is a particular case of the binomial negative INGARCH process introduced in Zhu (2011), and such that

$$
P(X_t = k|\underline{X}_{t-1}) = \frac{1}{1+\lambda_t}\left(1-\frac{1}{1+\lambda_t}\right)^k, \quad k \in \mathbb{N}_0,
$$

where $\lambda_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{j=1}^q \beta_j \lambda_{t-j}$. It is also obtained in Gonçalves et al (2015), and noted $\mathrm{NB}\left(1, \frac{1}{1+\lambda_t}\right)$, considering, conditionally to $\underline{X}_{t-1}$,

$$
X_t = Y_{t,1} + Y_{t,2} + \ldots + Y_{t,N_t}
$$

with $Y_{t,1}, Y_{t,2}$, ..., independent and identically distributed random variables with logarithmic distribution with parameter $\frac{1}{1+\lambda_t}$, independent of the random variable N_t which follows a Poisson law with parameter $\ln(1+\lambda_t)$. The characteristic function of the compounding variables, $Y_{t,j}$, is given by

$$\varphi_t(u) = \frac{\ln\left(1 - \frac{\lambda_t}{1+\lambda_t}\exp(iu)\right)}{\ln\left(\frac{1}{1+\lambda_t}\right)}.$$

which is a $\underline{X}_{t-1}-$ measurable and dependent on t function.

First and second order stationarity conditions and the study of the autocorrelation function of the geometric INGARCH process may be found in Zhu (2011) and Gonçalves et al (2015). In particular, we have

$$\begin{aligned} E\left(X_t|\underline{X}_{t-1}\right) &= \frac{1-\frac{1}{1+\lambda_t}}{\frac{1}{1+\lambda_t}} = \lambda_t \\ V\left(X_t|\underline{X}_{t-1}\right) &= \frac{1-\frac{1}{1+\lambda_t}}{\left(\frac{1}{1+\lambda_t}\right)^2} = \lambda_t\left(1+\lambda_t\right), \end{aligned}$$

This model allows integer-valued time series with overdispersion, as we deduce that $V(X_t) > E(X_t)$.

4.1.2. *The model for* $Z_t = X_{1,t} - X_{2,t}$.

Let us consider that the processes $(X_{1,t}, t \in \mathbb{Z})$ and $(X_{2,t}, t \in \mathbb{Z})$ are $\mathrm{NB}\left(1, \frac{1}{1+M_t}\right)$ and $\mathrm{NB}\left(1, \frac{1}{1+\widetilde{M}_t}\right)$, respectively, with

$$M_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j X_{1,t-j} + \sum_{k=1}^{q} \beta_k M_{t-k}$$

and

$$\widetilde{M_t} = \widetilde{\alpha}_0 + \sum_{j=1}^{\widetilde{p}} \widetilde{\alpha}_j X_{2,t-j} + \sum_{k=1}^{\widetilde{q}} \widetilde{\beta}_k \widetilde{M}_{t-k}.$$

The probability function of the conditional law of the process difference $Z_t = X_{1,t} - X_{2,t}$, $t \in \mathbb{Z}$, is given by

$$P\left(Z_t = k \mid \underline{Z}_{t-1}\right) = \begin{cases} \dfrac{1}{1+M_t+\widetilde{M}_t}\left(\dfrac{M_t}{1+M_t}\right)^k, & k \in \mathbb{N}_0 \\ \dfrac{1}{1+M_t+\widetilde{M}_t}\left(\dfrac{\widetilde{M}_t}{1+\widetilde{M}_t}\right)^{-k}, & -k \in \mathbb{N}. \end{cases}$$
$$= \frac{1}{1+M_t+\widetilde{M}_t}\left(\frac{M_t}{1+M_t}\right)^{k1_{\mathbb{N}_0}(k)}\left(\frac{\widetilde{M}_t}{1+\widetilde{M}_t}\right)^{-k1_{\mathbb{N}}(-k)}.$$

4.1.3. *Conditional maximum likelihood estimation*.

To describe the maximum likelihood approach to estimate the parameter vector

$$\begin{aligned}\Theta &= \left(\alpha_0, ..., \alpha_p, \beta_1, ..., \beta_q, \widetilde{\alpha}_0, ..., \widetilde{\alpha}_{\widetilde{p}}, \widetilde{\beta}_1, ..., \widetilde{\beta}_{\widetilde{q}}\right)^T \\ &= \left(\theta_1, ..., \theta_{p+1}, \theta_{p+2}, ..., \theta_{p+q+1}, \theta_{p+q+2}, ..., \theta_p, \theta_{p+1}, ..., \theta_{p+\widetilde{p}+q+\widetilde{q}+2}\right)^T\end{aligned}$$

of a stochastic process Z following a signed geometric INGARCH model, we note that the conditional likelihood function associated to n observations $Z_1, ..., Z_n$ conditionally to the initial values is

$$L(\Theta) = \prod_{t=1}^{n} \tfrac{1}{1+M_t+\widetilde{M}_t}\left(\tfrac{M_t}{1+M_t}\right)^{Z_t 1_{\mathbb{N}_0}(Z_t)}\left(\tfrac{\widetilde{M}_t}{1+\widetilde{M}_t}\right)^{-Z_t 1_{\mathbb{N}}(-Z_t)}$$

The log-likelihood function is given by

$\mathcal{L}(\Theta) = \log L(\Theta) = \sum_{t=1}^{n} l_t(\Theta)$

with

$l_t(\Theta) = -\log\left(1 + M_t + \widetilde{M}_t\right) + Z_t\left[1_{\mathbb{N}_0}(Z_t)\log\left(\tfrac{M_t}{1+M_t}\right) - 1_{\mathbb{N}}(-Z_t)\log\left(\tfrac{\widetilde{M}_t}{1+\widetilde{M}_t}\right)\right].$

To estimate the true value of Θ, Θ_0, it is natural to maximize $\mathcal{L}(\Theta)$ but, as the estimates has no closed form, numerical optimization methods have to be used.

In order to estimate the asymptotic covariance matrix of the conditional maximum likelihood estimator, $\widehat{\Theta}$, namely $[n\mathcal{I}(\Theta_0)]^{-1}$ where $\mathcal{I}(\Theta_0)$ is the information matrix evaluated at Θ_0, let us begin by considering the first derivatives of l_t in order to the first parameters θ_i, $i = 1, ..., p+q+1$, namely,

$$\begin{aligned}
\frac{\partial l_t}{\partial \theta_i} &= -\frac{1}{1+M_t+\widetilde{M}_t}\frac{\partial M_t}{\partial \theta_i} + Z_t 1_{\mathbb{N}_0}(Z_t)\left[\frac{1}{M_t} - \frac{1}{1+M_t}\right]\frac{\partial M_t}{\partial \theta_i} \\
&= \left[-\frac{1}{1+M_t+\widetilde{M}_t} + Z_t 1_{\mathbb{N}_0}(Z_t)\frac{1}{M_t(1+M_t)}\right]\frac{\partial M_t}{\partial \theta_i} \qquad (11)
\end{aligned}$$

and the second derivatives

$$\begin{aligned}
\frac{\partial^2 l_t}{\partial \theta_i \partial \theta_j} &= \left[-\frac{1}{1+M_t+\widetilde{M}_t} + \frac{Z_t 1_{\mathbb{N}_0}(Z_t)}{M_t(1+M_t)}\right]\frac{\partial^2 M_t}{\partial \theta_i \partial \theta_j} + \\
&\quad + \left[\frac{1}{\left(1+M_t+\widetilde{M}_t\right)^2}\frac{\partial M_t}{\partial \theta_j} + Z_t 1_{\mathbb{N}_0}(Z_t)\frac{-\frac{\partial M_t}{\partial \theta_j}(1+M_t) - M_t\frac{\partial M_t}{\partial \theta_j}}{M_t^2(1+M_t)^2}\right]\frac{\partial M_t}{\partial \theta_i} \\
&= \left[-\frac{1}{1+M_t+\widetilde{M}_t} + \frac{Z_t 1_{\mathbb{N}_0}(Z_t)}{M_t(1+M_t)}\right]\frac{\partial^2 M_t}{\partial \theta_i \partial \theta_j} + \\
&\quad + \left[\frac{1}{\left(1+M_t+\widetilde{M}_t\right)^2} - \frac{Z_t 1_{\mathbb{N}_0}(Z_t)(1+2M_t)}{M_t^2(1+M_t)^2}\right]\frac{\partial M_t}{\partial \theta_i}\frac{\partial M_t}{\partial \theta_j}, \qquad (12)
\end{aligned}$$

for $1 \leq i, j \leq p+q+1$. Moreover,

$$\begin{aligned}
\frac{\partial M_t}{\partial \alpha_0} &= 1 + \sum_{k=1}^{q}\beta_k \frac{\partial M_{t-k}}{\partial \alpha_0}; \\
\frac{\partial M_t}{\partial \alpha_i} &= Z_{t-i} + \sum_{k=1}^{q}\beta_k \frac{\partial M_{t-k}}{\partial \alpha_i}, i = 1, ..., p, \\
\frac{\partial M_t}{\partial \beta_j} &= M_{t-j} + \sum_{k=1}^{q}\beta_k \frac{\partial M_{t-k}}{\partial \beta_j}, j = 1, ..., q.
\end{aligned}$$

Taking expectations in both sides of the equation (12) we obtain

$$\begin{aligned}
E\left(\frac{\partial^2 l_t}{\partial \theta_i \partial \theta_j}|\underline{Z}_{t-1}\right) &= \left(-\frac{1}{1+M_t+\widetilde{M}_t} + \frac{E\left(Z_t 1_{\mathbb{N}_0}(Z_t)|\underline{Z}_{t-1}\right)}{M_t(1+M_t)}\right)\frac{\partial^2 M_t}{\partial \theta_i \partial \theta_j} + \\
&\quad + \left(\frac{1}{\left(1+M_t+\widetilde{M}_t\right)^2} - \frac{(1+2M_t)E\left(Z_t 1_{\mathbb{N}_0}(Z_t)|\underline{Z}_{t-1}\right)}{M_t^2(1+M_t)^2}\right)\frac{\partial M_t}{\partial \theta_i}\frac{\partial M_t}{\partial \theta_j}.
\end{aligned}$$

But from

$$E\left(Z_t 1_{\mathbb{N}_0}(Z_t)\,|\underline{Z}_{t-1}\right)=\frac{1}{1+M_t+\widetilde{M}_t}\sum_{k=0}^{+\infty}k\left(\frac{M_t}{1+M_t}\right)^k=\frac{1+M_t}{1+M_t+\widetilde{M}_t}\frac{1-\frac{1}{1+M_t}}{\frac{1}{1+M_t}}=\frac{1+M_t}{1+M_t+\widetilde{M}_t}M_t$$

we deduce that

$$E(\frac{\partial^2 l_t}{\partial\theta_i\partial\theta_j}|\underline{Z}_{t-1})=\left(\frac{1}{\left(1+M_t+\widetilde{M}_t\right)^2}-\frac{(1+2M_t)}{M_t^2\left(1+M_t\right)^2}E\left(Z_t 1_{\mathbb{N}_0}(Z_t)\,|\underline{Z}_{t-1}\right)\right)\frac{\partial M_t}{\partial\theta_i}\frac{\partial M_t}{\partial\theta_j}$$

$$=\left(\frac{1}{\left(1+M_t+\widetilde{M}_t\right)^2}-\frac{(1+2M_t)}{M_t\left(1+M_t\right)\left(1+M_t+\widetilde{M}_t\right)}\right)\frac{\partial M_t}{\partial\theta_i}\frac{\partial M_t}{\partial\theta_j}.$$

In an analogous way, from (11) we get

$$E(\frac{\partial l_t}{\partial\theta_i}\frac{\partial l_t}{\partial\theta_j}|\underline{Z}_{t-1})=E\left((-\frac{1}{1+M_t+\widetilde{M}_t}+Z_t 1_{\mathbb{N}_0}(Z_t)\frac{1}{M_t\left(1+M_t\right)})^2|\underline{Z}_{t-1}\right)\frac{\partial M_t}{\partial\theta_i}\frac{\partial M_t}{\partial\theta_j}.$$

Taking into account that $E\left(Z_t^2 1_{\mathbb{N}_0}(Z_t)\,|\underline{Z}_{t-1}\right)=\dfrac{M_t\left(1+M_t\right)\left(1+2M_t\right)}{1+M_t+\widetilde{M}_t}$ we obtain

$$E\left(\frac{\partial l_t}{\partial\theta_i}\frac{\partial l_t}{\partial\theta_j}|\underline{Z}_{t-1}\right)$$

$$=\ (\frac{1}{\left(1+M_t+\widetilde{M}_t\right)^2}-\frac{2E\left(Z_t 1_{\mathbb{N}_0}(Z_t)\,|\underline{Z}_{t-1}\right)}{M_t\left(1+M_t\right)\left(1+M_t+\widetilde{M}_t\right)}+\frac{E\left(Z_t^2 1_{\mathbb{N}_0}(Z_t)\,|\underline{Z}_{t-1}\right)}{M_t^2\left(1+M_t\right)^2})\frac{\partial M_t}{\partial\theta_i}\frac{\partial M_t}{\partial\theta_j}$$

$$=\ \left[-\frac{1}{\left(1+M_t+\widetilde{M}_t\right)^2}+\frac{1+2M_t}{M_t\left(1+M_t\right)\left(1+M_t+\widetilde{M}_t\right)}\right]\frac{\partial M_t}{\partial\theta_i}\frac{\partial M_t}{\partial\theta_j},$$

deducing that $-E\left(\dfrac{\partial^2 l_t}{\partial\theta_i\partial\theta_j}|\underline{Z}_{t-1}\right)=E\left(\dfrac{\partial l_t}{\partial\theta_i}\dfrac{\partial l_t}{\partial\theta_j}|\underline{Z}_{t-1}\right),i,j=1,...,p+q+1.$

We obtain analogously the first and second derivatives of l_t in order to the other parameters θ_i, $i=p+q+2,...,p+\widetilde{p}+q+\widetilde{q}+2$, namely

$$\frac{\partial l_t}{\partial\theta_i}=\left[-\frac{1}{1+M_t+\widetilde{M}_t}-\frac{Z_t 1_{\mathbb{N}}(-Z_t)}{\widetilde{M}_t\left(1+\widetilde{M}_t\right)}\right]\frac{\partial\widetilde{M}_t}{\partial\theta_i}\qquad(13)$$

$$\begin{aligned}\frac{\partial^2 l_t}{\partial\theta_i\partial\theta_j} &= \left[-\frac{1}{1+M_t+\widetilde{M}_t}-\frac{Z_t 1_{\mathbb{N}}(-Z_t)}{\widetilde{M}_t\left(1+\widetilde{M}_t\right)}\right]\frac{\partial^2\widetilde{M}_t}{\partial\theta_i\partial\theta_j}+\\ &+\left[\frac{1}{\left(1+M_t+\widetilde{M}_t\right)^2}+\frac{Z_t 1_{\mathbb{N}}(-Z_t)\left(1+2\widetilde{M}_t\right)}{\widetilde{M}_t^2\left(1+\widetilde{M}_t\right)^2}\right]\frac{\partial\widetilde{M}_t}{\partial\theta_i}\frac{\partial\widetilde{M}_t}{\partial\theta_j}\end{aligned} \tag{14}$$

Proceeding as above we deduce that the usual information matrix equality follows

$$-E\left(\frac{\partial^2 l_t}{\partial\theta_i\partial\theta_j}|\underline{Z}_{t-1}\right)=E\left(\frac{\partial l_t}{\partial\theta_i}\frac{\partial l_t}{\partial\theta_j}|\underline{Z}_{t-1}\right), i,j=1,...,p+\widetilde{p}+q+\widetilde{q}+2.$$

Asymptotic standard errors of the conditional maximum likelihood estimators of Θ can be computed from the matrix $\frac{1}{n}\left(\widehat{D}_n\widehat{S}_n\widehat{D}_n\right)^{-1}$ where

$$\widehat{S}_n=\frac{1}{n}\sum_{t=1}^{n}\frac{\partial l_t}{\partial\Theta}\frac{\partial l_t}{\partial\Theta^T} \text{ and } \widehat{D}_n=-\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2 l_t}{\partial\Theta\partial\Theta^T}.$$

4.2. Simulation study.

A simulation study was carried out to evaluate the finite sample performance of the CML estimators.

Table 1 presents the sample means and the standard deviations for the CML estimates of $\alpha_0, \alpha_1, \beta_1, \widetilde{\alpha}_0, \widetilde{\alpha}_1$ and $\widetilde{\beta}_1$ for the signed geometric model with $M_t = \alpha_0 + \alpha_1 X_{1,t-1} + \beta_1 M_{t-1}$ and $\widetilde{M}_t = \widetilde{\alpha}_0 + \widetilde{\alpha}_1 X_{2,t-1} + \widetilde{\beta}_1 \widetilde{M}_{t-1}$. These estimates were obtained considering different sample sizes, namely $n = 1000, 4000, 10000$ and $\alpha_0 = \widetilde{\alpha}_0 = 1.2, \alpha_1 = \widetilde{\alpha}_1 = \beta_1 = \widetilde{\beta}_1 = 0.2$.

We generated a sample of the signed geometric model of size n and, for this sample, we obtained its CML estimates following the previous theoretical approach. We repeated this procedure 100 times and the mean values of the estimates, along with the standard deviations in parenthesis, are presented in Table 1.

These simulations show that, as expected, the estimates of the six parameters seem to converge to the corresponding true parameter values as the sample size

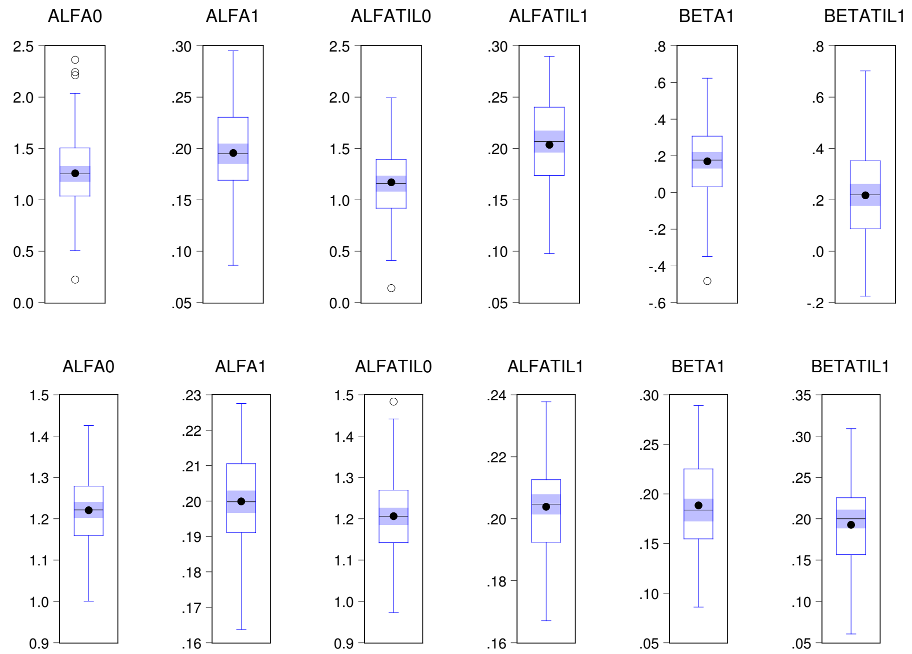


FIGURE 1. Box-plots of the CML estimates for all parameters, with $n = 1000$ and 10000.

increases. Further, the standard deviations of the estimates decrease when the sample size increases.

Table 1. CML estimates for the signed geometric model with $\alpha_0 = \widetilde{\alpha}_0 = 1.2, \alpha_1 = \widetilde{\alpha}_1 = \beta_1 = \widetilde{\beta}_1 = 0.2$, for sample sizes $n = 1000, 4000$ and 10000.

	$E_{est}\left(\widehat{\alpha_0}\right)$	$E_{est}\left(\widehat{\alpha_1}\right)$	$E_{est}\left(\widehat{\beta_1}\right)$	$E_{est}\left(\widehat{\widetilde{\alpha}_0}\right)$	$E_{est}\left(\widehat{\widetilde{\alpha}_1}\right)$	$E_{est}\left(\widehat{\widetilde{\beta}_1}\right)$
$n = 1000$	1.260643 (0.405936)	0.195888 (0.042905)	1.17689 (0.211527)	1.173820 (0.364874)	0.203735 (0.044899)	0.218863 (0.187787)
$n = 4000$	1.193396 (0.172529)	0.1968091 (0.023841)	0.203823 (0.083176)	0.176754 (0.175265)	0.203517 (0.024281)	0.210494 (0.090910)
$n = 10000$	1.220624 (0.089579)	0.199929 (0.014106)	0.188485 (0.047199)	1.206462 (0.103817)	0.203864 (0.014282)	0.192802 (0.053779)

Figure 1 displays the Box-plots of the CML estimates for all parameters for $n = 1000$ and 10000. We observe, in all cases, a strong concentration in the vicinity of the true value of the parameter and, paying attention to the scale of the plot, we notice that this concentration increases with n.

The Q-Q plots corresponding to $n = 1000$ and 10000 are presented in Figure 2. We note that the empirical quantiles approach the Gaussian distribution ones when n increases, for all the parameters in study. Moreover, the Jarque-Bera statistics and p-values presented in Table 2 for $n = 10000$ show the clear compatibility of the CML estimates of the parameters with the Gaussian distribution.

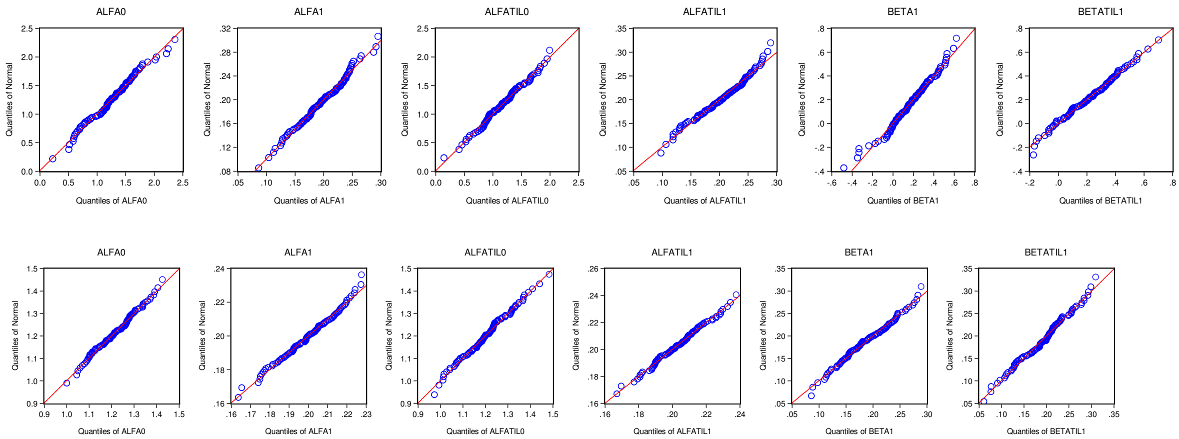


FIGURE 2. Q-Q plots of the CML estimates for $n = 1000$ and 10000.

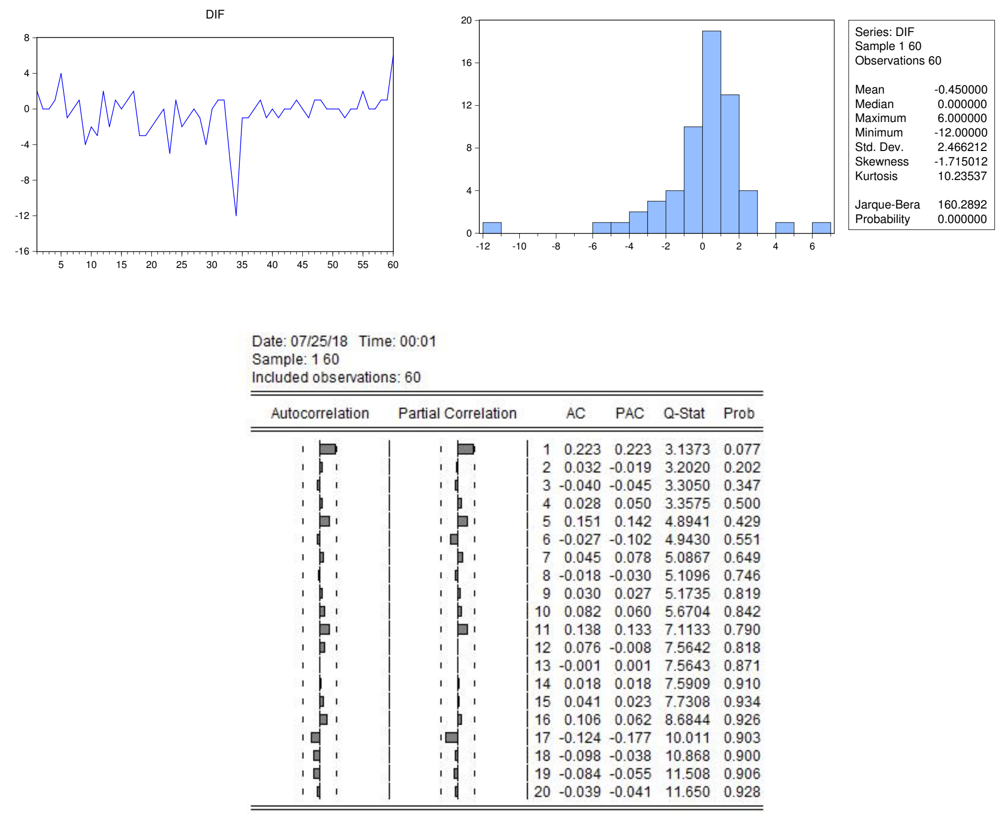
Table 2. CML estimates and Gaussian distribution for $n = 10000$.

	α_0	α_1	β_1	$\widetilde{\alpha}_0$	$\widetilde{\alpha}_1$	$\widetilde{\beta}_1$
Jarque-Bera	0.961911	1.658994	1.869042	0.188014	0.118713	0.90713
Probability	0.618193	0.436269	0.392774	0.910276	0.942371	0.635361

4.3. Applications.

4.3.1. *Poliomyelitis cases in USA.* We apply the proposed estimation methodology to the polio data discussed in Zeger (1988) and Zhu (2011), among others. The data consists of monthly counts of poliomyelitis cases recorded in the United States from 1970 to 1983 by the Centres for Disease Control. We consider two subseries of this data as spaced as possible, namely the monthly counts of polio cases from 1970 to 1974 and from 1979 to 1983 and we study the series that is the difference between the most recent values and those more distant (60 observations).

Figure 3 presents this series, its descriptive summaries and empirical autocorrelation and partial autocorrelation values. The empirical mean and standard



	AC	PAC	Q-Stat	Prob
1	0.223	0.223	3.1373	0.077
2	0.032	-0.019	3.2020	0.202
3	-0.040	-0.045	3.3050	0.347
4	0.028	0.050	3.3575	0.500
5	0.151	0.142	4.8941	0.429
6	-0.027	-0.102	4.9430	0.551
7	0.045	0.078	5.0867	0.649
8	-0.018	-0.030	5.1096	0.746
9	0.030	0.027	5.1735	0.819
10	0.082	0.060	5.6704	0.842
11	0.138	0.133	7.1133	0.790
12	0.076	-0.008	7.5642	0.818
13	-0.001	0.001	7.5643	0.871
14	0.018	0.018	7.5909	0.910
15	0.041	0.023	7.7308	0.934
16	0.106	0.062	8.6844	0.926
17	-0.124	-0.177	10.011	0.903
18	-0.098	-0.038	10.868	0.900
19	-0.084	-0.055	11.508	0.906
20	-0.039	-0.041	11.650	0.928

FIGURE 3. Difference series: plot, descriptive summaries and autocorrelation and partial autocorrelation values.

deviation of the data are -0.45 and 2.466 respectively. We conclude that there was, on average, progress in that 10-year interval in the direction of polio eradication in United States. The data is overdispersed, the autocorrelation of order one is 0.223 and the autocorrelations of higher order are not significant, which allows inferring an order 1 dependence although not very strong.

In order to model the data we consider a signed geometric INARCH model with $p = \widetilde{p} = 1$ with parameters $\alpha_0, \alpha_1, \widetilde{\alpha}_0, \widetilde{\alpha}_1$. The conditional maximum likelihood estimates from this fitting are summarised in Table 3 (0.509345, 0.293369, 0.519798 and 0.522338, respectively) leading to an estimated model that is first order stationary.

The fitted conditional mean is given in Figure 4 and we note that it closely follows the values and trend of the observed series. The resulting Pearson residual is defined by

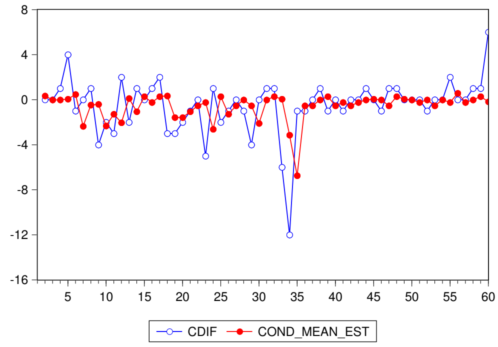


FIGURE 4. Difference series and fitted conditional mean from the signed geometric INGARCH model.

$$r_t = \frac{X_t - \left(\widehat{M}_t - \widehat{\widehat{M}}_t\right)}{\sqrt{\widehat{M}_t\left(1+\widehat{M}_t\right) + \widehat{\widehat{M}}_t\left(1+\widehat{\widehat{M}}_t\right)}}$$

where $\widehat{M}_t = \widehat{\alpha}_0 + \widehat{\alpha}_1 X_{1,t-1}$ and $\widehat{\widehat{M}}_t = \widehat{\widehat{\alpha}}_0 + \widehat{\widehat{\alpha}}_1 X_{2,t-1}$. The residual analysis is shown in Figure 5 and there is no evidence of any correlation within the residuals. The Jarque-Bera statistics implies the normality of the Pearson residuals at the 0.01 significance level and this fact is also suggested by the kernel density estimation and normal Q-Q plots for the Pearson residuals.

Table 3. Signed geometric INARCH model parameter estimation

LogL: LOGL03
Sample: 2 60
Date: 07/24/18 Time: 11:32
Included observations: 59
Evaluation order: By equation
Convergence achieved after 19 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.509345	0.314045	1.621885	0.1048
C(2)	0.293369	0.200107	1.466060	0.1426
C(3)	0.519798	0.262866	1.977421	0.0480
C(4)	0.522338	0.225187	2.319573	0.0204

Log likelihood	-118.0980	Akaike info criterion	4.138915
Avg. log likelihood	-2.001661	Schwarz criterion	4.279765
Number of Coefs.	4	Hannan-Quinn criter.	4.193897

4.3.2. *Olympic Medals won by Swiss and Dutch athletes.* Our second application consists of the number of Olympic medals won by Switzerland (S) and Netherlands (N) as displayed in https://demos.telerik.com/ aspnet-ajax/sample-applications/olympic-games/.

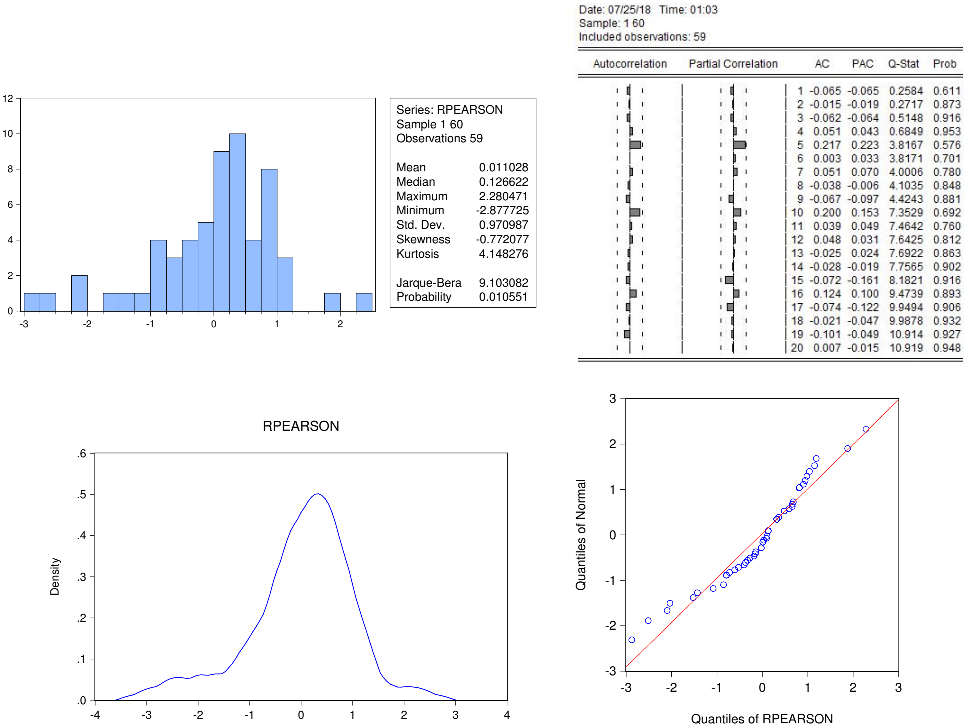


FIGURE 5. Pearson residuals: descriptive summaries, autocorrelation and partial autocorrelation values, kernel density estimation and Gaussian Q-Q plot.

Figure 6 presents the plot of the S-N difference series, its descriptive summaries and the empirical autocorrelation and partial autocorrelation values. The empirical mean and variance are -1.935 and 76.74, respectively. A better performance, on average, of Dutch athlets is observed. The data is overdispersed and the autocorrelations of order greater than two are not significant.

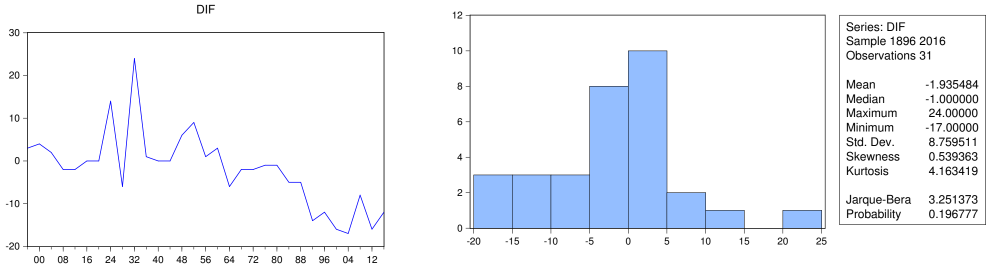
The conditional maximum likelihood estimates after fitting a signed geometric INARCH model with parameters $\alpha_0, \alpha_1, \widetilde{\alpha}_0, \widetilde{\alpha}_2$ are 1.64603, 0.306622, 2.145151 and 0.476046, respectively (Table 4).

Table 4. Signed geometric INARCH model parameter $\alpha_0, \alpha_1, \widetilde{\alpha}_0, \widetilde{\alpha}_2$ estimation

LogL: LOGL01
Sample: 1904 2016
Date: 09/19/18 Time: 10:45
Included observations: 29
Evaluation order: By equation
Convergence achieved after 18 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	1.646030	0.805580	2.043285	0.0410
C(2)	0.306622	0.136229	2.250782	0.0244
C(3)	2.145151	1.260393	1.701969	0.0888
C(4)	0.476046	0.284534	1.673073	0.0943

Log likelihood	-96.28852	Akaike info criterion	6.916450
Avg. log likelihood	-3.320294	Schwarz criterion	7.105042
Number of Coefs.	4	Hannan-Quinn criter.	6.975515



Date: 09/16/18 Time: 19:16
Sample: 1896 2016
Included observations: 31

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.421	0.421	6.0371	0.014
		2	0.583	0.493	18.028	0.000
		3	0.361	0.046	22.777	0.000
		4	0.323	-0.076	26.723	0.000
		5	0.311	0.094	30.541	0.000
		6	0.185	-0.053	31.939	0.000
		7	0.240	0.040	34.389	0.000
		8	0.028	-0.180	34.424	0.000
		9	0.074	-0.068	34.675	0.000
		10	-0.096	-0.120	35.129	0.000
		11	-0.099	-0.082	35.629	0.000
		12	-0.077	0.093	35.950	0.000
		13	-0.142	0.009	37.101	0.000
		14	-0.138	-0.100	38.251	0.000
		15	-0.276	-0.180	43.121	0.000
		16	-0.247	-0.077	47.286	0.000

FIGURE 6. Difference series: plot, descriptive summaries and autocorrelation and partial autocorrelation values.

Despite the significance of all the estimated parameters we decided, in view of the small significance of the order 2 autocorrelation, to fit the series by a a signed geometric INARCH model with parameters $\alpha_0, \alpha_1, \widetilde{\alpha}_0, \widetilde{\alpha}_1$. The corresponding estimates are now 1.374252, 0.368949, 0.81656 and 0.622359, respectively (Table 5). The values of the Akaike and Schwarz criteria lead us to retain this second modeling.

Table 5. Signed geometric INARCH model parameter $\alpha_0, \alpha_1, \widetilde{\alpha}_0, \widetilde{\alpha}_1$ estimation

LogL: LOGL01
Method: Maximum Likelihood (Newton-Raphson / Marquardt steps)
Date: 07/28/18 Time: 22:26
Sample: 1900 2016
Included observations: 30
Evaluation order: By equation
Convergence achieved after 6 iterations
Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	1.374252	0.619727	2.217510	0.0266
C(2)	0.368949	0.152283	2.422782	0.0154
C(3)	0.816560	0.784855	1.040397	0.2982
C(4)	0.622359	0.319859	1.945727	0.0517

Log likelihood	-97.12910	Akaike info criterion	6.741940
Avg. log likelihood	-3.237637	Schwarz criterion	6.928766
Number of Coefs.	4	Hannan-Quinn criter.	6.801707

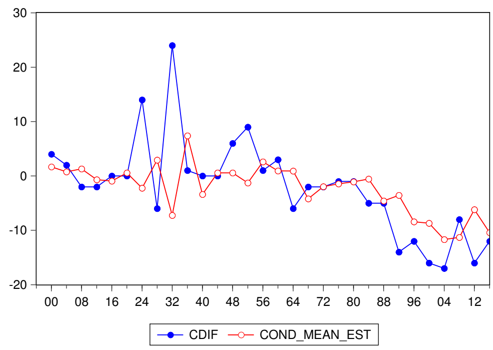


FIGURE 7. Difference series and fitted conditional mean from the signed geometric INGARCH model

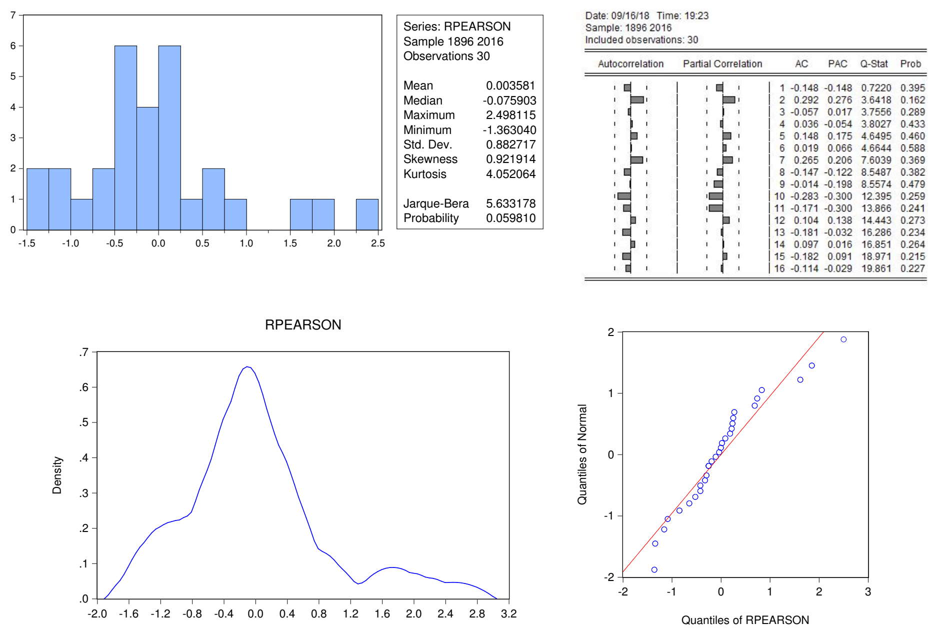


FIGURE 8. Pearson residuals: descriptive summaries, autocorrelation and partial autocorrelation values, kernel density estimation and Gaussian Q-Q plot.

So, considering this first order model we observe that the fitted conditional mean accompanies the dynamics of the observed series (Figure 7). The resulting Pearson residual analysis given in Figure 8 shows that there is no evidence of correlation within the residuals and that there is compatibility with Gaussian distribution at 0.01 and 0.05 significance levels. We observe that the analysis of the Pearson residuals for the second order model, firstly considered, led to worst results confirming the decision based on the criteria.

5. Conclusion

Time series of counts appear in a large variety of contexts like in studies of the incidence of a certain disease in a country, number of daily transactions on a financial market or number of accidents in a town. This kind of time series often reveals overdispersion and conditional heteroscedasticity and the large family of integer-valued CP-INGARCH models has wide potential to describe and capture these characteristics.

With the aim of modeling the difference of two count time series, we propose in this paper a bivariate model defined by two independent CP-INGARCH processes. A $\mathbb{Z}$-valued counting process is then defined as the difference between the two marginal processes. The signed integer-valued process defined by the Skellam distribution, which is constructed as differences in pairs of Poisson counts, is included in this study if the counts series are independent.

Since the Poissonian models are not the most adequated to model overdispersed series we concentrated our study in the geometric models, a particular case of the NB-INARCH ones. The probabilistic and statistical study here developed shows that this family of $\mathbb{Z}$-valued models may be useful in applications where the analysis of the difference of count time series is relevant.

Although we have privileged the difference of the two marginal processes, we note that any other measurable function of these processes may be considered and emphasize that the main probabilistic properties of such models have already been here stated.

Acknowledgements

This work was partially supported by the Centre for Mathematics of the University of Coimbra - UID/MAT/00324/2013 funded by the Portuguese Government trough FCT/MEC and co-funded by the European Regional Development Fund through the Partnership Agreement PT2020.

References

[1] Abramowitz M, Stegun I (eds.). 1965. *Handbook of Mathematical Functions*, New York: Dover Publications, Inc.

[2] Durrett R. 2010. *Probability: Theory and Examples*, Cambridge University Press (4th edn).

[3] Ferland R, Latour A, Oraichi D. 2006. *Integer-valued GARCH process.*, J. Time Ser. Analysis, 27, 923–42.

[4] Fokianos K, Rahbek A, Tjøtheim D. 2009. *Poisson autoregression*, Journal of the American Statistical Association, 104, 1430–39.

[5] Gonçalves E, Mendes-Lopes N, Silva F. 2015. *Infinitely divisible distributions in integer-valued GARCH models*, J. Time Ser. Analysis, 36, 503-527.

[6] Gonçalves E, Mendes-Lopes N, Silva F. 2016. *Zero-inflated compound Poisson distributions in integer-valued GARCH models*, Statistics, 50, 558-578.

[7] Halmos P. 1974. *Measure Theory*, New York: Springer-Verlag.

[8] Karlis D, Ntzoufras I. 2006. *Bayesian analysis of the differences of count data*, Statistics in Medicine, 25, 1885–1905.

[9] Karlis D, Ntzoufras I. 2009. *Bayesian modelling of football outcomes: using the Skellam distribution for the goal difference*, IMAJ of Management Mathematics, 20, 133–145.

[10] Kim HY, Park Y. 2008. *A non-stationary integer-valued autoregressive model*, Statistical Papers, 49, 485–502.

[11] Koopman SJ, Lit R, Lucas A. 2014. *The Dynamic Skellam Model with Applications*, Tinbergen Institute Discussion Paper 14-032/IV/DSF73, Available at SSRN: https://ssrn.com/abstract=2406867 or http://dx.doi.org/10.2139/ssrn.2406867.

[12] Scotto MG, Weiß CH, Gouveia S. 2015.*Thinning-based models in the analysis of integer-valued time series: A review*, Statistical Modelling, 15(6): 590–618.

[13] Skellam JG. 1946.*The frequency distribution of the difference between two Poisson variates belonging to different populations*, J. of the Royal Statistical Society, A, 109 (3), 296-.

[14] Weiß CH. 2009. *Modelling time series of counts with overdispersion*, Statistical Methods and Applications, 18, 507–519.

[15] Xu H-Y, Xie M, Goh TN., Fu X. 2012. *A model for integer-valued time series with conditional overdispersion*,Computational Statistics and Data Analysis, 56, 4229–4242.

[16] Zeger SL. 1988.*A regression model for time series of counts*, Biometrika, 75, 621-629.

[17] Zhu F. 2011. *A negative binomial integer-valued GARCH model*, J. Time Ser. Anal. 32, 54–67.

[18] Zhu F. 2012. *Modelling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models*,Journal of Mathematical Analysis and Applications, 389 1, 58–71.

[19] Zhu F, Li Q, Wang D. 2010.*A mixture integer-valued ARCH model*, Journal of Stat. Plann. and Inf., 140, 2025–2036.

E. Gonçalves
CMUC, Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal
E-mail address: esmerald@mat.uc.pt

N. Mendes-Lopes
CMUC, Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal
E-mail address: nazare@mat.uc.pt