A new approach to integer-valued time series modeling:
The Neyman type-A INGARCH model

E. Gonçalves[1], N. Mendes-Lopes, F. Silva

CMUC, Department of Mathematics, University of Coimbra

Abstract. The aim of this paper is to develop a probabilistic study on a wide class of conditionally heteroscedastic models recently introduced in literature, the Compound Poisson INGARCH processes ([8]). This class includes, in particular, some well-known models like the Poisson INGARCH of Ferland, Latour and Oraichi ([5]) or the negative binomial and the generalized Poisson INGARCH introduced by Zhu in 2011 and 2012, respectively.

We analyse the existence, within this class, of a strictly and weakly stationary solution, as well as its ergodicity. For a new particular model of that class, the Neyman type-A INGARCH one, we derive the autocorrelation function, analyse the existence of higher-order moments and obtain the explicit form of their first four cumulants, from which the corresponding skewness and kurtosis are deduced.

Keywords: integer-valued time series, GARCH model, infinitely divisible discrete probability laws, compound Poisson distributions, Neyman type-A distribution.

1 Introduction

Integer-valued time series, quite common in various contexts and scientific fields like medicine, economics, finance or epidemiology, are not well reproduced by real-valued stochastic processes, namely when dealing with low dimension samples. In fact, in this case, disregarding the nature of the data leads, in general, to senseless results as the asymptotic behavior of the corresponding statistical parameters or distributions is not available. Among the integer-valued models presented in literature in last decade, we highlight the INGARCH one, ([5]), analogous to GARCH models ([2]) but with Poisson deviates, that is, the process at time t given its past follows a Poisson distribution. This definition, which is a natural way of introducing conditional heteroscedasticity in integer time series, motivated the proposal and study of other INGARCH models replacing the distribution of deviates by new discrete laws like negative binomial ([14]) or generalized Poisson ([15]). With the same goal, we introduce an integer-valued process with general infinitely divisible deviates enlarging and unifying this INGARCH family ([8]).

With this new definition a wide set of probability distributions for deviates is considered including, in particular, those related to the models referred above. Moreover, we may identify this set with the family of the probability distributions of a Poissonian random sum of independent variables with discrete distribution ([11]) and so, more than to present a model for which the conditional distribution is compound Poisson, we can define it clearly as a counting process. So, this general approach leads us to a lot of recent contributions on integer-valued time series modeling and several new interesting particular cases, as the Neyman type-A INGARCH model, may be constructed. We point out that the Neyman type-A distribution is widely used in describing populations under the influence of contagion, what is particularly interesting when we know that integer-valued time series are well-adapted to model bacteriologic diseases ([5]).

A sufficient condition for the existence of a strict and weakly stationary and ergodic solution of these models is presented in Section 2. In Section 3 we focus on the Neyman type-A INGARCH $(1,1)$ model obtaining its autocorrelation function, establishing a necessary and sufficient condition for the existence of higher-order moments and giving explicitly the first four cumulants

[1]Corresponding author: Esmeralda Gonçalves E-mail: esmerald@mat.uc.pt CMUC, Department of Mathematics, University of Coimbra, Apartado 3008, EC Santa Cruz, 3001-501 Coimbra, Portugal.

of the process. As a consequence, the corresponding skewness and kurtosis are deduced. We conclude with some remarks and future developments.

2 The compound Poisson INGARCH model

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

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

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.

In the previous definition, when $q = 1$ and $\beta_1 = 0$ the CP-INGARCH(p, q) model is simply denoted by CP-INARCH(p).

The conditional mean and variance of X_t are given by

$$E(X_t|\underline{X}_{t-1}) = \lambda_t, \quad V(X_t|\underline{X}_{t-1}) = -i\frac{\varphi_t''(0)}{\varphi_t'(0)}\lambda_t,$$

considering $(\varphi_t, t \in \mathbb{Z})$ derivable at zero up to order 2, to assure the variance existence.

This model includes several models recently studied in the literature and also a wide class of new processes where the characteristic function φ_t may be a random characteristic function or a deterministic one. Namely, as it is established in [8], in the second case it is included the Poisson INGARCH model, introduced by Ferland, Latour and Oraichi [5], the Generalized Poisson INGARCH due to Zhu [15], or the negative binomial DINARCH model (Xu et al. [13]); otherwise the first case includes, in particular, the negative binomial INGARCH model, (Zhu [14]).

To illustrate the large class of new particular models satisfying this framework, let us recall that since 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 with

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

where N_t follows a Poisson law with parameter $\lambda_t^* = i\ \lambda_t/\varphi_t'(0)$, and $X_{t,1}$,..., 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.

Using that generation some new examples may be introduced.

Example 2.1 ***(a)*** *Let us consider independent random variables $(X_{t,j}, t \in \mathbb{Z})$ following the same discrete distribution which has finite mean, constant parameters and support contained in $\mathbb{N}_0$. If N_t is a random variable independent of $X_{t,j}$ and following a Poisson law with parameter $\frac{\lambda_t}{E(X_{t,j})}$, the process $X_t = \sum_{j=1}^{N_t} X_{t,j}$ follows the model (1). For instance, we may consider the Poisson distribution with parameter $\theta > 0$ for $X_{t,j}$ and the parameter $\lambda_t^* = \frac{\lambda_t}{\theta}$ in the Poisson law of N_t.*

In this case, we obtain as conditional distribution the Neyman type-A one and the model will be denoted by NTA-INGARCH (p, q).

(b) *If* $(X_{t,j}, t \in \mathbb{Z})$ *are independent random variables following the binomial distribution with parameters* $r \in \mathbb{N}$ *and* $e^{-|t|}$*, that is,* $\varphi_t(u) = \left(e^{iu-|t|} + 1 - e^{-|t|}\right)^r$*,* $u \in \mathbb{R}$*,* $t \in \mathbb{Z}$*, and* N_t *is an independent of* $X_{t,j}$ *random variable following* $\mathcal{P}(\frac{\lambda_t}{re^{-|t|}})$ *then* X *satisfies the model* (1).

We note that these processes are first-order stationary if and only if $\sum_{j=1}^{p} \alpha_j + \sum_{k=1}^{q} \beta_k < 1$. Moreover, under this condition the processes (X_t) and (λ_t) are both first-order stationary and we have

$$E(X_t) = E(\lambda_t) = \mu = \frac{\alpha_0}{1 - \sum_{j=1}^{p} \alpha_j - \sum_{k=1}^{q} \beta_k}.$$

Now, we focus on the subclass of CP-INGARCH (p, q) models for which the characteristic functions φ_t are deterministic and independent of t. This particular case includes a wide class of models not studied in literature as the referred NTA-INGARCH model.

In the following theorem we resume a sufficient condition under which there exists a strictly stationary and ergodic CP-INGARCH (p, q) process.

Theorem 2.1 *If* $\sum_{j=1}^{p} \alpha_j + \sum_{k=1}^{q} \beta_k < 1$*, then there is a strictly stationary and ergodic process that satisfies the model* (1)*. Moreover, its first two moments are finite.*

Proof. 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. As $\beta_1 + ... + \beta_q < 1$, the roots of $B(z) = 0$ lie outside the unit circle and we can write

$$\lambda_t = B^{-1}(L) \left[\alpha_0 + A(L) X_t\right] = \psi_0 + \sum_{j=1}^{\infty} \psi_j X_{t-j},$$

where $\psi_0 = \alpha_0 B^{-1}(1)$ and ψ_j is the coefficient of z^j in the MacLaurin expansion of $H(z) = \frac{A(z)}{B(z)}$.

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

$$\Phi_{U_t}(u) = \exp\left\{\psi_0 \frac{i}{\varphi'(0)} \left[\varphi(u) - 1\right]\right\}.$$

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

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

where $(\psi_j, j \in \mathbb{N})$ is the sequence of the above coefficients. Note that $E(U_t) = \psi_0$, $E(Z_{t,k,j}) = \psi_k$ 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 defined 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

$$X_t^{(n)} = \begin{cases} 0, & n < 0 \\ U_t, & n = 0 \\ U_t + \sum_{k=1}^{n} \sum_{j=1}^{X_{t-k}^{(n-k)}} Z_{t-k,k,j}, & n > 0 \end{cases}, \tag{3}$$

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

Using the thinning operation (see, e.g., [12]), the sequence is rewritten as

$$X_t^{(n)} = U_t + \sum_{k=1}^{n} \psi_k^{(t-k)} \circ X_{t-k}^{(n-k)}, \quad n > 0,$$

where the notation $(\psi_k^{(\tau)}\circ)$ means that the sequence of random variables of common mean ψ_k involved in the thinning operation corresponds to time τ.

The expectation and the variance of $X_t^{(n)}$ are well defined, because $X_t^{(n)}$ is a finite sum of independent compound Poisson random variables with characteristic functions φ derivable at zero up to order 2. $E[X_t^{(n)}]$ does not depend on t, it just depends on n and will be denoted by μ_n. In fact, the result is trivial for $n<0$. For $n=0$ we obtain $E(X_t^{(0)})=E(U_t)=\psi_0$, which is also independent of t. Let us consider now, as induction hypothesis, that for an arbitrarily fixed value of t and until $n>0$, $E[X_t^{(n)}]$ is independent of t. Therefore,

$$E\left(X_t^{(n+1)}\right)=\psi_0+\sum_{k=1}^{n+1}\psi_k E\left(X_{t-k}^{(n+1-k)}\right)=g\left(E\left(X_{t-n-1}^{(0)}\right),...,E\left(X_{t-1}^{(n)}\right)\right).$$

Using (3), the property of the thinning operation $E(\phi\circ W)=\phi E(W)$ and the fact that $\mu_j=0$ if $j<0$, we have, for $n>0$,

$$\mu_n=\psi_0+\sum_{k=1}^{\infty}\psi_k\mu_{n-k}=B^{-1}(L)\left[\alpha_0+A(L)\mu_n\right]\Leftrightarrow K(L)\mu_n=\alpha_0,$$

where $K(L)=B(L)-A(L)$. So, $\{(X_t^{(n)},t\in\mathbb{Z}),\ n\in\mathbb{Z}\}$ is a sequence of first-order stationary processes because the characteristic polynomial $K(z)$ has all its roots outside the unit circle, according to the hypothesis on the parameters of the model.

Using arguments similar to that of [5] and [8] we know that the sequence $\{(X_t^{(n)},t\in\mathbb{Z}),\ n\in\mathbb{Z}\}$ has an almost sure limit, $X^*=(X_t^*,t\in\mathbb{Z})$, and it is a strictly stationary process for each n. Then it is easily seen that X^* is a strictly stationary process.

Moreover, $(X_t^{(n)})$ is a sequence of ergodic processes, because it is a measurable function of the sequence of independent and identically distributed random variables $\{(U_t,\mathcal{Z}_{t,j}),t\in\mathbb{Z},j\in\mathbb{N}\}$ (see, e.g., [4]). Taking into account that (X_t^*) is the almost sure limit of a sequence of measurable functions, so a measurable one, then (X_t^*) is ergodic.

From the stationarity and Beppo Lévi's theorem

$$\mu=\lim_{n\to\infty}\mu_n=\lim_{n\to\infty}E\left(X_t^{(n)}\right)=E\left(X_t^*\right),$$

and we conclude that the first moment of X_t^* is finite.

For the second moment, we note that

$$E\left[\left(X_t^{(n)}\right)^2\right]\le\frac{E(X_t^*)\left(2\psi_0-i\frac{\varphi''(0)}{\varphi'(0)}\right)\sum_{k=1}^{\infty}\psi_k+E(U_t^2)}{1-\left(\sum_{k=1}^{\infty}\psi_k\right)^2}=C;$$

so, using the Lebesgue's dominated convergence theorem we conclude that $E(X_t^2)\le C$.

Finally, the almost sure limit of the sequence $(X_t^{(n)})$ is a solution of the model since

$$\Phi_{X_t^*|\underline{X}_{t-1}^*}(u)\overset{(1)}{=}\lim_{n\to+\infty}\Phi_n(u)\overset{(2)}{=}\exp\left\{i\frac{\lambda_t}{\varphi'(0)}\left[\varphi_t(u)-1\right]\right\},\quad u\in\mathbb{R},$$

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

$$r_t^{(n)}=U_t+\sum_{k=1}^{n}\sum_{j=1}^{X_{t-k}^*}Z_{t-k,k,j}.$$

In fact, the equality (1) follows from Paul Lévy theorem since, similarly to section 2.6 of [5], we prove that for a fixed t, the sequence $Y_t^{(n)}=r_t^{(n)}-X_t^{(n)}$ converges in mean to zero, when $n\to\infty$.

So $Y_t^{(n)}$ and $X_t^* - X_t^{(n)}$ converge in probability to zero and as $X_t^* - r_t^{(n)} = (X_t^* - X_t^{(n)}) - Y_t^{(n)}$, we conclude that the sequence $r_t^{(n)}$ converges in probability to X_t^* and therefore $r_t^{(n)}|\underline{X}_{t-1}^*$ converges in law to $X_t^*|\underline{X}_{t-1}^*$. Now, let us obtain Φ_n. Conditionally to $\underline{X}_{t-1}^*$, we have

$$\Phi_n(u) = \exp\left\{\left(\psi_0 + \sum_{k=1}^{n} \psi_k X_{t-k}^*\right) \frac{i}{\varphi'(0)} \left[\varphi_t(u) - 1\right]\right\},$$

and so, when $n \to \infty$, we conclude equality (2). □

From the proof of this theorem we may conclude that if $\sum_{k=1}^{q} \beta_k < 1$ we obtain a CP-INARCH (∞) representation of the process. Moreover, we can still deduce the weak stationarity of the process solution as it is a L^2 strictly stationary process.

3 The Neyman type-A integer-valued GARCH$(1,1)$ model

A random variable X follows a Neyman type-A distribution with parameters $\lambda > 0$ and $\theta > 0$, denoted by $Neyman(\lambda, \theta)$, if its probability mass function is

$$P(X = x) = \frac{\exp\left\{-\lambda + \lambda e^{-\theta}\right\} \theta^x}{x!} \sum_{j=0}^{x} S(x,j) \lambda^j e^{-j\theta}, \quad x = 0, 1, ...,$$

where the coefficient $S(x, j)$ represents the Stirling number of the second kind. We note that $S(0, 0) = 1$, $S(x, 0) = 0$, for $x \neq 0$, $S(x, 1) = S(x, x) = 1$, $S(x, j) = 0$ if $j > x$ and these numbers satisfy the recurrence relation $S(x, j) = S(x-1, j-1) + jS(x-1, j)$ (for details, see [1]).

The probability generating function of the Neyman type-A distribution is

$$g(z) = \exp\left\{\lambda\left(e^{\theta(z-1)} - 1\right)\right\},$$

and $E(X) = \lambda\theta$, $V(X) = \lambda\theta(1+\theta)$. This distribution is an example of a compound Poisson law where its compounding distribution is also a Poisson law. The Neyman type-A distribution has been found useful to explain phenomena in various areas of application such as biology, ecology or in the context of bus driver accidents (see [10], Section 9.6.1, and references therein).

Based on this, let us consider the CP-INGARCH model where the conditional law of X_t is the Neyman type-A distribution with parameters $(\frac{\lambda_t}{\theta}, \theta)$, as stated in **(a)** of Example 2.1.

In the following, we focus on the NTA-INGARCH model with $p = q = 1$, *i.e.*,

$$\Phi(u) = \exp\left\{\lambda_t/\theta\left[\exp\left(\theta(e^{iu} - 1)\right) - 1\right]\right\}, \quad \lambda_t = \alpha_0 + \alpha_1 X_{t-1} + \beta_1 \lambda_{t-1}.$$

3.1 Stationarity, ergodicity and moments existence

The stationarity and ergodicity of the NTA-INGARCH $(1,1)$ process are assured, from Theorem 2.1, under the condition $\alpha_1 + \beta_1 < 1$. In the following we show that this coefficients condition is necessary and sufficient to establish the existence of all the moments of that process.

Theorem 3.1 *The moments of a NTA-INGARCH* $(1,1)$ *process are all finite if, and only if,* $\alpha_1 + \beta_1 < 1$.

Proof. According to [9], since $X_t|\underline{X}_{t-1}$ is a compound Poisson random variable where the compounding distribution is the Poisson law with parameter θ, its mth moment is given by

$$E[X_t^m|\underline{X}_{t-1}] = \sum_{r=0}^{m} \frac{1}{r!} \frac{\lambda_t^r}{\theta^r} \sum_{j=0}^{r} \binom{r}{j} \frac{(-1)^{r-j}}{i^m} \left(\varphi^j\right)^{(m)}(0),$$

where $\left(\varphi^j\right)^{(m)}(u)$ represents the mth derivative of $\varphi^j(u) = \prod_{k=1}^{j} \varphi(u) = \exp\left(j\theta\left(e^{iu}-1\right)\right)$, with $\varphi^j(u) = 0$ if $j = 0$. So, the following expression holds

$$\left(\varphi^j\right)^{(m)}(u) = i^m \sum_{k=1}^{m} S(m,k)(j\theta)^k \exp\left(j\theta\left(e^{iu}-1\right)+kiu\right), \quad u \in \mathbb{R}. \tag{4}$$

In fact, for $m = 1$, we have

$$\left(\varphi^j(u)\right)' = ij\theta \exp\left(j\theta\left(e^{iu}-1\right)+iu\right)$$

and, by induction,

$$\left(\varphi^j\right)^{(m+1)}(u) = \frac{d}{du}\left[i^m \sum_{k=1}^{m} S(m,k)(j\theta)^k \exp\left(j\theta\left(e^{iu}-1\right)+kiu\right)\right]$$

$$= i^m \sum_{k=1}^{m} S(m,k)(j\theta)^k \left(ij\theta e^{iu}+ik\right) \exp\left(j\theta\left(e^{iu}-1\right)+kiu\right)$$

$$= i^{m+1} \sum_{k=1}^{m} S(m,k)\left[(j\theta)^{k+1} \exp\left(j\theta\left(e^{iu}-1\right)+(k+1)iu\right) + k(j\theta)^k \exp\left(j\theta\left(e^{iu}-1\right)+kiu\right)\right]$$

$$= i^{m+1} j\theta \exp\left(j\theta\left(e^{iu}-1\right)+iu\right) + i^{m+1}(j\theta)^{m+1} \exp\left(j\theta\left(e^{iu}-1\right)+i(m+1)u\right)$$

$$+ i^{m+1} \sum_{k=2}^{m} (j\theta)^k [S(m-1,k)+kS(m,k)] \exp\left(j\theta\left(e^{iu}-1\right)+kiu\right)$$

$$= i^{m+1} \sum_{k=1}^{m+1} S(m+1,k)(j\theta)^k \exp\left(j\theta\left(e^{iu}-1\right)+kiu\right).$$

So

$$\left(\varphi^j\right)^{(m)}(0) = i^m \sum_{k=1}^{m} S(m,k)(j\theta)^k.$$

Thus,

$$E[X_t^m] = \sum_{r=0}^{m} \frac{1}{r!} \sum_{j=0}^{r} \binom{r}{j} \frac{(-1)^{r-j}}{\theta^r} \sum_{k=1}^{m} S(m,k)(j\theta)^k \; E[\lambda_t^r], \tag{5}$$

and

$$\lambda_t^r = (\alpha_0 + \alpha_1 X_{t-1} + \beta_1 \lambda_{t-1})^r = \sum_{n=0}^{r} \binom{r}{n} \alpha_0^{r-n} \sum_{l=0}^{n} \binom{n}{l} \alpha_1^l \beta_1^{n-l} X_{t-1}^l \lambda_{t-1}^{n-l}.$$

Using the fact that λ_{t-1}^{n-l} is $\underline{X}_{t-2}$-measurable we deduce that

$$E[\lambda_t^r | \underline{X}_{t-2}] = \sum_{n=0}^{r} \binom{r}{n} \alpha_0^{r-n} \sum_{l=0}^{n} \binom{n}{l} \alpha_1^l \beta_1^{n-l} \lambda_{t-1}^{n-l} \; E[X_{t-1}^l | \underline{X}_{t-2}]$$

$$= \sum_{n=0}^{r} \binom{r}{n} \alpha_0^{r-n} \sum_{l=0}^{n} \binom{n}{l} \sum_{v=0}^{l} \frac{\alpha_1^l \beta_1^{n-l}}{v! \; \theta^v} \sum_{k=1}^{l} \sum_{j=0}^{v} (-1)^{v-j} S(l,k) \binom{v}{j} (j\theta)^k \lambda_{t-1}^{v+n-l}, \tag{6}$$

assuming that $\sum_1^0 = 1$. Let $\Lambda_t = (\lambda_t^m, ..., \lambda_t)'$. In the algebraic expression of $E[\lambda_t^r | \underline{X}_{t-2}]$, for $r = 1, ..., m$, all the powers of λ_{t-1} are $\leq r$. Therefore, a constant vector $\mathbf{d}$ and an upper triangular matrix $\mathbf{D} = (d_{ij})$, $i, j = 1, ..., m$, exist such that the following equation is satisfied:

$$\begin{bmatrix} E[\lambda_t^m | \underline{X}_{t-2}] \\ \vdots \\ E[\lambda_t^2 | \underline{X}_{t-2}] \\ E[\lambda_t | \underline{X}_{t-2}] \end{bmatrix} = \begin{bmatrix} \alpha_0^m \\ \vdots \\ \alpha_0^2 \\ \alpha_0 \end{bmatrix} + \begin{bmatrix} (\alpha_1+\beta_1)^m & \cdots & d_{1,m-1} & d_{1m} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & (\alpha_1+\beta_1)^2 & d_{m-1,m} \\ 0 & \cdots & 0 & \alpha_1+\beta_1 \end{bmatrix} \begin{bmatrix} \lambda_{t-1}^m \\ \vdots \\ \lambda_{t-1}^2 \\ \lambda_{t-1} \end{bmatrix}$$

$$\Leftrightarrow E[\Lambda_t|\underline{X}_{t-2}] = \mathbf{d} + \mathbf{D}\Lambda_{t-1}.$$

Let us prove that the diagonal entries of the matrix $\mathbf{D}$ are those given above.

The ith diagonal entry of the matrix $\mathbf{D}$ corresponds to the case where in equation (6), we consider $r = m - i + 1$. Thus, to obtain the coefficient of λ_{t-1}^{m-i+1}, we look at the terms corresponding to $n = m - i + 1$ and $l = v$. Then,

$$d_{ii} = \sum_{l=0}^{m-i+1} \left(\begin{array}{c} m-i+1 \\ l \end{array} \right) \frac{\alpha_1^l \beta_1^{m-i+1-l}}{l!\theta^l} \sum_{k=1}^{l} \sum_{j=0}^{l} (-1)^{l-j} S(l,k) \left(\begin{array}{c} l \\ j \end{array} \right) (j\theta)^k$$

and using the closed form of the Stirling numbers of the second kind ([1]) we deduce

$$d_{ii} = \sum_{l=0}^{m-i+1} \left(\begin{array}{c} m-i+1 \\ l \end{array} \right) \alpha_1^l \beta_1^{m-i+1-l} \sum_{k=1}^{l} S(l,k) S(k,l) \theta^{k-l} = (\alpha_1 + \beta_1)^{m-i+1},$$

because $S(k,l) \neq 0$ only when $k = l$.

So, we conclude that the eigenvalues of $\mathbf{D}$ (which coincide with its diagonal entries because it is a triangular matrix) are inside the unit circle if, and only if, $\alpha_1 + \beta_1 < 1$. Using this fact and proceeding as in Proposition 6 of [5], we can write

$$E[\Lambda_t|\underline{X}_{t-k}] = (I_m - \mathbf{D})^{-1}(I_m - \mathbf{D}^{k-1})\mathbf{d} + \mathbf{D}^{k-1}\Lambda_{t-(k-1)},$$

where I_m is the identity matrix of order m. Consequently, since $\mathbf{D}^{k-1} \to 0$ when $k \to \infty$, we have $E[\Lambda_t] = \lim_{k\to\infty} E[\Lambda_t|\underline{X}_{t-k}] = (I_m - \mathbf{D})^{-1}\mathbf{d}$, and then from (5) all the moments of X_t of order $\leq m$ are finite and independent of t. $\square$

In the next theorem we establish the autocovariance function of a NTA-INGARCH$(1,1)$ process.

Theorem 3.2 *Let X be a NTA-INGARCH$(1,1)$ process such that $\alpha_1 + \beta_1 < 1$. The autocovariance function of X is given by*

$$\Gamma(k) = \left\{ \begin{array}{ll} \frac{\mu(1+\theta)\left(1-(\alpha_1+\beta_1)^2+\alpha_1^2\right)}{1-(\alpha_1+\beta_1)^2}, & k = 0 \\ \frac{\alpha_1\mu(1+\theta)(1-\beta_1(\alpha_1+\beta_1))(\alpha_1+\beta_1)^{k-1}}{1-(\alpha_1+\beta_1)^2}, & k > 0 \end{array} \right. ,$$

with $\mu = \frac{\alpha_0}{1-(\alpha_1+\beta_1)}$.

***Proof*.** Let us start by noting that

$$\begin{array}{ll} E(X_{t-j}\lambda_{t-k}) = E\left[E(X_{t-j}|\underline{X}_{t-j-1})\lambda_{t-k}\right] = E(\lambda_{t-j}\lambda_{t-k}), & k \geq j, \\ E(X_{t-j}\lambda_{t-k}) = E\left[X_{t-j}E(X_{t-k}|\underline{X}_{t-k-1})\right] = E(X_{t-j}X_{t-k}), & k < j, \end{array}$$

$$\Leftrightarrow E(X_{t-j}\lambda_{t-k}) = \left\{ \begin{array}{ll} E(\lambda_{t-j}\lambda_{t-k}), & \text{se } k \geq j \\ E(X_{t-j}X_{t-k}), & \text{se } k < j \end{array} \right. .$$

Using the above expression we deduce

$$E(X_t X_{t-k}) = E\left[E(X_t|\underline{X}_{t-1})X_{t-k}\right] = E\left(\left[\alpha_0 + \alpha_1 X_{t-1} + \beta_1\lambda_{t-1}\right] X_{t-k}\right)$$

$$= \alpha_0\mu + \alpha_1 E(X_{t-1}X_{t-k}) + \beta_1 E(\lambda_{t-1}X_{t-k}) = \alpha_0\mu + (\alpha_1+\beta_1)E(X_{t-1}X_{t-k}), \quad k \geq 2,$$

$$E(\lambda_t\lambda_{t-k}) = E\left(\left[\alpha_0 + \alpha_1 X_{t-1} + \beta_1\lambda_{t-1}\right]\lambda_{t-k}\right)$$

$$= \alpha_0\mu + \alpha_1 E(X_{t-1}\lambda_{t-k}) + \beta_1 E(\lambda_{t-1}\lambda_{t-k}) = \alpha_0\mu + (\alpha_1+\beta_1)E(\lambda_{t-1}\lambda_{t-k}), \quad k \geq 1,$$

and then the autocovariances $\Gamma(k) = Cov(X_t, X_{t-k})$ and $\widetilde{\Gamma}(k) = Cov(\lambda_t, \lambda_{t-k})$ satisfy

$$\begin{aligned}\Gamma(k) &= (\alpha_1+\beta_1)\cdot\Gamma(k-1) = ... = (\alpha_1+\beta_1)^{k-1}\cdot\Gamma(1)\\ &= (\alpha_1+\beta_1)^{k-1}[\alpha_1\cdot\Gamma(0)+\beta_1\cdot\widetilde{\Gamma}(0)], \qquad k\geq 2,\\ \widetilde{\Gamma}(k) &= (\alpha_1+\beta_1)\cdot\widetilde{\Gamma}(k-1) = ... = (\alpha_1+\beta_1)^{k}\cdot\widetilde{\Gamma}(0), \qquad k\geq 1,\end{aligned}$$

using the equality $\mu - \alpha_0 = \mu(\alpha_1+\beta_1)$. As

$$\begin{aligned}E(\lambda_t^2) &= \alpha_0\mu + \alpha_1 E(X_{t-1}X_t) + \beta_1 E(\lambda_{t-1}\lambda_t) \quad \Rightarrow \widetilde{\Gamma}(0) = \alpha_1\cdot\Gamma(1) + \beta_1\cdot\widetilde{\Gamma}(1),\\ E(X_tX_{t-1}) &= \alpha_0\mu + \alpha_1 E(X_{t-1}^2) + \beta_1 E(\lambda_{t-1}^2) \quad \Rightarrow \Gamma(1) = \alpha_1\cdot\Gamma(0) + \beta_1\cdot\widetilde{\Gamma}(0),\end{aligned}$$

we conclude that

$$\widetilde{\Gamma}(0) = \alpha_1^2\cdot\Gamma(0) + \alpha_1\beta_1\cdot\widetilde{\Gamma}(0) + \beta_1(\alpha_1+\beta_1)\cdot\widetilde{\Gamma}(0) = \alpha_1^2\cdot\Gamma(0) + [(\alpha_1+\beta_1)^2 - \alpha_1^2]\cdot\widetilde{\Gamma}(0)$$

$$\Leftrightarrow \widetilde{\Gamma}(0) = V(\lambda_t) = \frac{\alpha_1^2\cdot\Gamma(0)}{1-(\alpha_1+\beta_1)^2+\alpha_1^2}.$$

Then, the autocovariance function of X, for $k \geq 1$, is given by

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

where

$$\begin{aligned}E(X_t^2) = (1+\theta)E(\lambda_t) + E(\lambda_t^2) &\Leftrightarrow \Gamma(0) = (1+\theta)\mu + \widetilde{\Gamma}(0)\\ &\Leftrightarrow \Gamma(0) = \frac{\mu(1+\theta)\left(1-(\alpha_1+\beta_1)^2+\alpha_1^2\right)}{1-(\alpha_1+\beta_1)^2}.\end{aligned}$$

□

We remark that the autocovariance function of the process (λ_t), $\widetilde{\Gamma}$, is also obtained from this theorem.

3.2 Skewness and Kurtosis

The knowledge and properties of a stochastic process moments are particularly useful in the description or estimation of its unknown probability distribution. Skewness and kurtosis are moment parameters who give relevant information on the symmetry and shape of the related distribution. In the following, we deduce the skewness and kurtosis of a NTA-INARCH (1) process, X, such that $\alpha_1 < 1$, beginning by providing closed-form expressions for their cumulants up to order 4. With this purpose, let us recall that if Φ_{X_t} denotes the characteristic function of X_t, its cumulant generating function is given by $\kappa_{X_t}(z) = \ln(\Phi_{X_t}(z))$, and the coefficient $\kappa_j(X_t)$ of the series expansion $\kappa_{X_t}(z) = \sum_{j=1}^{\infty}\kappa_j(X_t)\cdot(iz)^j/j!$ is referred to as the cumulant with $\kappa_j(X_t) = (-i)^j\kappa_{X_t}^{(j)}(0)$. To obtain the referred cumulants, let us consider the expression of the characteristic function of X_t, namely, for $z\in\mathbb{R}$,

$$\begin{aligned}\Phi_{X_t}(z) &= E\left(e^{izX_t}\right) = E\left[E\left(e^{izX_t}|\underline{X}_{t-1}\right)\right] = E\left[\exp\left\{\lambda_t/\theta\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right\}\right]\\ &= E\left[\exp\left\{\alpha_0/\theta\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right\}\exp\left\{\alpha_1X_{t-1}/\theta\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right\}\right]\\ &= \exp\left(\frac{\alpha_0}{\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right)\cdot\Phi_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right),\end{aligned}$$

hence, the cumulant generating function of X_t is given by

$$\kappa_{X_t}(z) = \frac{\alpha_0}{\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right] + \kappa_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right).$$

Taking derivatives on both sides, it follows that

$$\begin{aligned}\kappa'_{X_t}(z) &= i\alpha_0 \exp\left(\theta(e^{iz}-1)+iz\right)\\ &\quad + \alpha_1 \exp\left(\theta(e^{iz}-1)+iz\right)\cdot \kappa'_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right). \end{aligned} \tag{7}$$

Inserting $z=0$ into the previous equation, we obtain

$$\kappa'_{X_t}(0) = i\alpha_0 + \alpha_1\cdot\kappa'_{X_{t-1}}(0) \Rightarrow \kappa_1(X_t) = \frac{\alpha_0}{1-\alpha_1} = \mu.$$

Taking derivatives on both sides of the equation (7),

$$\begin{aligned}\kappa''_{X_t}(z) &= -\alpha_0\left(\theta e^{iz}+1\right)\exp\left(\theta(e^{iz}-1)+iz\right)\\ &\quad + i\alpha_1\left(\theta e^{iz}+1\right)\exp\left(\theta(e^{iz}-1)+iz\right)\cdot\kappa'_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right)\\ &\quad + \alpha_1^2\exp\left(2\theta(e^{iz}-1)+2iz\right)\cdot\kappa''_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right)\\ &= i\left(\theta e^{iz}+1\right)\cdot\kappa'_{X_{t-1}}(z)\\ &\quad + \alpha_1^2\exp\left(2\theta(e^{iz}-1)+2iz\right)\cdot\kappa''_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right), \end{aligned} \tag{8}$$

so, inserting $z=0$ in this equation, we get

$$\kappa''_{X_t}(0) = i(1+\theta)\cdot\kappa'_{X_t}(0) + \alpha_1^2\cdot\kappa''_{X_t}(0) \Rightarrow \kappa_2(X_t) = V(X_t) = \frac{1+\theta}{1-\alpha_1^2}\mu,$$

which coincides with the expression stated in Theorem 3.2.

Taking derivatives on both sides of the equation (8), it follows that

$$\begin{aligned}\kappa'''_{X_t}(z) &= -\theta e^{iz}\cdot\kappa'_{X_t}(z) + i\left(\theta e^{iz}+1\right)\cdot\kappa''_{X_t}(z)\\ &\quad + 2i\alpha_1^2\left(\theta e^{iz}+1\right)\exp\left(2\theta(e^{iz}-1)+2iz\right)\cdot\kappa''_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right)\\ &\quad +\alpha_1^3\exp\left(3\theta(e^{iz}-1)+3iz\right)\cdot\kappa'''_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right)\\ &= \left[2\left(\theta e^{iz}+1\right)^2-\theta e^{iz}\right]\cdot\kappa'_{X_t}(z) + 3i\left(\theta e^{iz}+1\right)\cdot\kappa''_{X_t}(z)\\ &\quad +\alpha_1^3\exp\left(3\theta(e^{iz}-1)+3iz\right)\cdot\kappa'''_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right), \end{aligned} \tag{9}$$

and inserting $z=0$ in the equation we obtain

$$\begin{aligned}\kappa'''_{X_t}(0) &= \left[2(\theta+1)^2-\theta\right]\cdot\kappa'_{X_t}(0) + 3i(\theta+1)\cdot\kappa''_{X_t}(0) + \alpha_1^3\cdot\kappa'''_{X_{t-1}}(0)\\ &\Rightarrow \kappa_3(X_t) = \frac{\theta(1-\alpha_1^2)+(1+\theta)^2(1+2\alpha_1^2)}{(1-\alpha_1^2)(1-\alpha_1^3)}\mu.\end{aligned}$$

To obtain the fourth cumulant, let us take derivatives on both sides of the equation (9):

$$\begin{aligned}\kappa^{(iv)}_{X_t}(z) &= \left[4\left(\theta e^{iz}+1\right)-1\right]i\theta e^{iz}\cdot\kappa'_{X_t}(z) + \left(2\left(\theta e^{iz}+1\right)^2-\theta e^{iz}\right)\cdot\kappa''_{X_t}(z)\\ &\quad -3\theta e^{iz}\cdot\kappa''_{X_t}(z) + 3i\left(\theta e^{iz}+1\right)\cdot\kappa'''_{X_t}(z)\\ &\quad + \alpha_1^3\left(3i\theta e^{iz}+3i\right)\exp\left(3\theta(e^{iz}-1)+3iz\right)\cdot\kappa'''_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right)\end{aligned}$$

$$+\alpha_1^4 \exp\left(4\theta(e^{iz}-1)+4iz\right)\cdot\kappa^{(iv)}_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right)$$

$$=\left[\left(\theta e^{iz}+1\right)\left[7i\theta e^{iz}-6i\left(\theta e^{iz}+1\right)^2\right]-i\theta e^{iz}\right]\cdot\kappa'_{X_t}(z)$$

$$+\left[11\left(\theta e^{iz}+1\right)^2-4\theta e^{iz}\right]\cdot\kappa''_{X_t}(z)+6i\left(\theta e^{iz}+1\right)\cdot\kappa'''_{X_t}(z)$$

$$+\alpha_1^4 \exp\left(4\theta(e^{iz}-1)+4iz\right)\cdot\kappa^{(iv)}_{X_{t-1}}\left(\frac{\alpha_1}{i\theta}\left[\exp\left(\theta(e^{iz}-1)\right)-1\right]\right),$$

which, inserting $z=0$, implies that

$$\kappa^{(iv)}_{X_t}(0)=i\left[6\theta+7\theta^2-6(1+\theta)^3\right]\cdot\kappa'_{X_t}(0)+\left[11(1+\theta)^2-4\theta\right]\cdot\kappa''_{X_t}(0)$$

$$+6i(1+\theta)\cdot\kappa'''_{X_t}(0)+\alpha_1^4\cdot\kappa^{(iv)}_{X_{t-1}}(0)$$

$$\Rightarrow \kappa_4(X_t)=\frac{(7\alpha_1^2+6\alpha_1^3+5\alpha_1^5)(1+\theta)^3+(4\alpha_1^2+6\alpha_1^3-10\alpha_1^5)(\theta+\theta^2)}{(1-\alpha_1^2)(1-\alpha_1^3)(1-\alpha_1^4)}\mu$$

$$+\frac{(1+\theta)^3+3\theta^2+4\theta}{1-\alpha_1^4}\mu.$$

Using the relationship between cumulants and centered moments, the following result is easily established.

Theorem 3.3 *The skewness and kurtosis of a NTA-INARCH* (1) *process* X *such that* $\alpha_1<1$ *are given by*

$$S_{X_t}=\frac{\theta(1-\alpha_1^2)+(1+\theta)^2(1+2\alpha_1^2)}{(1+\theta)(1+\alpha_1+\alpha_1^2)}\sqrt{\frac{1+\alpha_1}{\alpha_0(1+\theta)}};$$

$$K_{X_t}=3+\frac{(1-\alpha_1)(1-\alpha_1^2)\left((1+\theta)^3+3\theta^2+4\theta\right)}{\alpha_0(1+\alpha_1^2)(1+\theta)^2}+\frac{(7\alpha_1^2+6\alpha_1^3+5\alpha_1^5)(1+\theta)}{\alpha_0(1+\alpha_1+\alpha_1^2)(1+\alpha_1^2)}$$

$$+\frac{\theta(4\alpha_1^2+6\alpha_1^3-10\alpha_1^5)}{\alpha_0(1+\alpha_1+\alpha_1^2)(1+\alpha_1^2)(1+\theta)}.$$

Finally, from those expressions, we conclude that X is an asymmetric, around the mean, and leptokurtic process.

In Figure 1 the trajectory and the basic descriptives of 1000 observations of a NTA-INARCH(1) process are presented from which is evident the closeness of the theoretical values $S_{X_t}\simeq 0,1004$ and $K_{X_t}\simeq 3,557$, according to the above formulas, and the empirical ones.

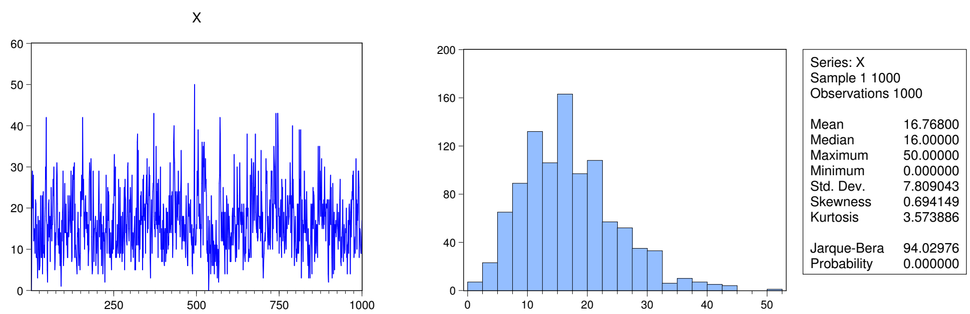

Figure 1: Trajectory and descriptives of NTA-INARCH(1) model: $\alpha_0=10$, $\alpha_1=0.4$, $\theta=2$.

4 Conclusion

Compound Poisson INGARCH processes have been recently introduced in literature ([8]) unifying and enlarging studies developed by Ferland, Latour and Oraichi (2006), or Zhu (2011, 2012), among others, to model the conditional volatility of integer-valued time series, analogously to what is considered in the GARCH processes of Bollerslev (1986). In fact, to consider the general compound Poisson law as the conditional distribution of the INGARCH process allows us to include in the same framework almost of those particular models and also to characterize clearly the counting process underlying the model; in consequence, an infinity of new models may be considered.

Taking into consideration the importance of the Neyman type-A (NTA) distribution in the description of populations under the influence of contagion, accidents, rare diseases, among others, and the relevance of considering integer-valued models when we deal with low dimension samples, we present here the main probabilistic developments of the NTA-INGARCH model. We stress that under very simple conditions on the model coefficients we assure the stationarity, weak and strict, and the ergodicity of the model as well as their moments existence. In consequence, we deduce the process skewness and kurtosis after calculating their first four cumulants. The probabilistic study developed is particularly relevant in future statistical studies as those related to process distribution estimation ([7]) or in the evalution of other features of this leptokurtic model like the Taylor property ([6]). Applications to real data are naturally the final purpose of such a study.

Acknowledgements. This work is supported by the Centro de Matemática da Universidade de Coimbra (funded by the European Regional Development Fund through the program COMPETE and by the Portuguese Government through the FCT - Fundação para a Ciência e Tecnologia under the projects PEst-C/MAT/UI0324/2011). The work of the third author was supported by a grant from the FCT with reference SFRH/BD/85336/2012.

References

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

[2] Bollerslev, T.: Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* **37**, 307–327 (1986).

[3] Consul, P. C., Famoye, F.: Lagrangian Probability Distributions. Birkhäuser, Boston (2006).

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

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

[6] Gonçalves, E., Leite, J., Mendes Lopes, N.: A mathematical approach to detect the Taylor property in TARCH processes. *Statistics and Probability Letters* **79**, 602–610 (2009).

[7] Gonçalves, E., Leite, J., Mendes Lopes, N.: On the finite dimensional laws of threshold GARCH processes. In *Recent Develoments in Modeling and Applications in Statistics*, Vichi, M., Oliveira, P.E., Temido, M.G., Henriques, C. (Eds.), Selected Papers of the Statistical Societies, Springer, 237–247 (2013).

[8] Gonçalves, E., Mendes Lopes, N., Silva, F.: Infinitely divisible distributions in integer valued GARCH models, *DMUC, Preprint 13-38* (2013).

[9] Grubbström, R. W., Tang, O.: The moments and central moments of a compound distribution. *European Journal of Operational Research* **170**, 106–119 (2006).

[10] Johnson, N.L., Kotz, S., Kemp, A.W.: Univariate discrete distributions. Wiley, New York, 3rd ed. (2005).

[11] Steutel, F. W., and van Harn, K.: Infinite divisibility of probability distributions on the real line. Marcel-Dekker, Inc, New York (2004).

[12] Weiß, C. H.: Thinning operations for modelling time series of counts - a survey. *Adv. Stat. Anal.* **92**, 319–341 (2008).

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

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

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