ON GENERALIZED INTEGER-VALUED GARCHX MODEL WITH STRUCTURAL CHANGES

Mohamed SADOUN

RECITS Laboratory, Operational Research Department, Faculty of Mathematics,
University of Sciences and Technology Houari Boumediene

ABSTRACT

We propose a generalized *INGARCH* model with structural changes including exogenous covariates (hereafter referred to as $GCP-INGARCHX$), which belongs to the observation-driven type models where the regime-switching is driven by certain faillure points occuring in the time. Necessary and sufficient conditions for the second-order stationarity, both in the mean and in the second-order moment, are established. The explicit formula of the mean and the second-order moment are then given under these second-order stationarity conditions. The autocovariance structure is studied while providing the closed-form of the piecewise autocorrelation function. The conditional least squares (*CLS*) and the conditional maximum likelihood (*CML*) estimators of the underlying parameters are obtained for both the cases that the break points are known or not. A simulation study and an application on real data set are provided to assess the performance of the model.

Keywords : Integer-valued process of counts, no-stationary *INGARCH* model, (*CLS*) estimators, (*CML*) estimators, piecewise autocorrelation function.
Mathematics Subject Classification 2010 : $62F12$, $62M10$

1. INTRODUCTION

In count time series analysis, Poisson distribution is frequently used and provides a classical framework (see Fokianos (2012)). However, due to the fact that many time series of counts encountered in practice exhibits the nonlinearity, the over and/or the under-dispersion, the multimodality, and the zeros-inflation, as well as the non-stationarity, a simple counts model with a regime-switching of change point-type including generalized and more flexible distributions is then needed to capture these characteristics often found in real-life examples. We propose a generalized integer-valued *GARCH* model with structural changes containing exogenous covariates (hereafter referred to as $GCP-INGARCHX$). Our formulated model is a generalized *INGARCHX* regime-switching model, where the regime-switching is driven by a certain latent random variable defined by pieces in time. Thus, we propose a more general threshold *INGARCHX* model, which includes several distributions for defining numerous *INGARCH* model cases, and where "X" refers to specific exogenous covariates. Our modeling approach generalizes some famous works in the time series of counts modeling, namely : the model introduced by Ferland *et al* (2006), the model introduced by Wang *et al* (2014), the nonlinear proposed model by Chen *et al* (2019), and the generalized mixture *INGARCH* model proposed by Mao et al (2019), and recently the Poisson and Negative Binomial *INGARCH* models proposed by Lee *et al* (2021) in their change-point test for the conditional mean of time series of counts.

2. MODEL FORMULATION AND MAIN ASSUMPTIONS

The definition of this generalized *INGARCHX* (P,Q,q) model with structural changes, is as follows :

$$\begin{cases} Z_t = \sum_{k=1}^{m+1} \mathbf{1}(S_t = k) Y_{S_t,t} \\ \mathbb{E}\left(Y_{S_t,t}|\mathscr{F}_{t-1}\right) = \lambda_{S_t,t}, \\ var\left(Y_{S_t,t}|\mathscr{F}_{t-1}\right) = \upsilon_{S_t,0}\lambda_{S_t,t} + \upsilon_{S_t,1}\lambda^2_{S_t,t} \\ \lambda_{S_t,t} = \alpha_{S_t,0} + \sum_{i=1}^{P} \alpha_{S_t,i} Z_{t-i} + \sum_{j=1}^{Q} \beta_{S_t,j} \lambda_{t-j} + \sum_{l=1}^{q} \gamma_{S_t,l} X_{t-l}. \end{cases} \tag{1}$$

where $\mathscr{F}_t$ indicates the information given up to time t. $\upsilon_{S_t,k} \geq 0, k = 0,1$, but not simultaneously equal to zero, $\alpha_{S_t,0} > 0$, $\alpha_{S_t,i} \geq 0$, $\beta_{S_t,j} \geq 0$, and $\gamma_{S_t,l} \geq 0$ for $i = 1,...,P$, $j = 1,...,Q$, and $l - 1,...,q$. $\mathbf{1}(.)$ denotes the indicator function. $\{S_t\}_{t\in\mathbb{Z}}$ is a sequence of independent and identically distributed random variables defined by pieces in time, which is defined as follows :

$$S_t = \begin{cases} 1, & \text{if } t \leq c_1, \\ 2, & \text{if } c_1 < t \leq c_2. \\ \vdots & \vdots \\ m & \text{if } c_{m-1} < t \leq c_m \\ m+1 & \text{if } t > c_m \end{cases}$$

where $(c_1, c_2, ..., c_m)'$ stand the vector of unknown break points. Furthermore, it is assumed that Z_{t-i} and S_t are independent for all t and $i > 0$. We note this model as $GCP-INGARCHX\,(P,Q,q)$. In this contribution we consider a simple version of the model (1) with $P = Q = 1$ and $d = 1$, namely the $GCP-INGARCHX\,(1,1,q)$ model where the change occures only on the link function $\lambda_{S_t,t}$ as follows :

$$\lambda_{S_t,t} = \alpha_{S_t,0} + \alpha_{S_t,1} Z_{t-1} + \beta_{S_t} \lambda_{t-1} + \sum_{l=1}^{q} \gamma_{S_t,l} X_{t-l}. \tag{2}$$

where $\alpha_{S_t,0} > 0$, $\alpha_{S_t,1} \geq 0$, $\beta_{S_t} \geq 0$, and $\gamma_{S_t,l} \geq 0$ for $l = 1,...,q$. $\{S_t\}_{t\in\mathbb{Z}}$ is still the same sequence of independent and identically distributed random variables, defined as in (1). The model (2) includes several models already studied in the literature. For example :

1. Wang et al (2014) proposed the self-excited thrshold Poisson autoregression model which is no other than the threshold *INGARCH* $(1,1)$ model by introducing Poisson distribution as conditional distribution of each component $Y_{S_t,t}$. This model result from the model (2) by puting $t = Z_{t-1}$, and by considering $\upsilon_{S_t,0} = 1$, $\upsilon_{S_t,1} = 0$, and $\gamma_{S_t,l} = 0$ for $l = 1,...,q$.
2. Chen et al (2019) considered a threshold specification of negative binomial integer-valued *GARCHX* $(1,1)$ model. This model is also including in model (2) by puting $t = Z_{t-1}$, and by setting $\upsilon_{S_t,0} = \upsilon_{S_t,1} = r$ such as r is a positif number.
3. Mao et al (2019) generalized the mixture *INGARCHX* model by proposing two different distributions as conditional distribution of each component $Y_{S_t,t}$, namely the Negative Binomial distribution and the Generalized Poisson distribution. This model result from model (2) by assuming the $\{S_t\}_{t\in\mathbb{Z}}$ is a sequence of independent and identically distributed random variables with some no spcify discrete distribution, where $\gamma_{S_t,l} = 0$ for $l = 1,...,q$.

It is worth mentioning that in a second-order point of view our proposed model namely $GCP-INGARCH$ model might be viewed like an *ARMA* model with structural breaks. This last transformation will facilitate us lot of computation regarding the probabilistic structure. Indeed, under suitable conditions, we show that all the moments exist and we give the mean and the autocovariance function of this generalized $CP-INGARCH$ process.

Through our formulation and specification of the conditional distribution concerning each component $Y_{S_t,t}$, we can obtain other *INGARCHX* models with structural breaks.

3. PARAMETERS ESTIMATION

For the problem of parameters estimation, we shall focus on the change-point negative binomial *INGARCHX* model and the change-point generalized Poisson *INGARCHX* model with following conditional distribution specifications :

$$Y_{S_t,t}|\mathscr{F}_{t-1} \rightsquigarrow \mathscr{N}\mathscr{B}\left(\lambda_{S_t,t}, \upsilon_{S_t}\right) \qquad (A.1)$$
$$Y_{S_t,t}|\mathscr{F}_{t-1} \rightsquigarrow \mathscr{G}\mathscr{P}\left((1-\phi_{S_t})\,\lambda_{S_t,t}, \phi_{S_t}\right) \qquad (A.2)\,.$$

Two methods of estimation will be considered in this paragraph, namely the *CLS* method and the *CML* method. Although the first one is less performant than the second one, it remains frequently used nowadays. Indeed, it is simple and it produces consistent estimators which can be used as initial values in sophisticated methods as the maximum likelihood estimation. We note that for the *CLS* method we used the *ARMA* presentation of our model to estimate the parameters.

Then, we considered two kinds of $GCP-INGARCHX\,(1,1)$ models based on different distributions as mentioned above.

For the first case of Negative Binomial *INGARCHX* model with $Y_{S_t,t}|\mathscr{F}_{t-1} \rightsquigarrow \mathscr{N}\mathscr{B}\left(\lambda_{S_t,t}, \upsilon_{S_t}\right)$, the conditional likelihood function of the n observations $Z_1,...,Z_n$ conditionally on the presample values, is given by

$$L(\Theta) = \prod_{t=1}^{n}\sum_{k=1}^{m+1}\left(\frac{\Gamma\left(Y_{k,t}+\lambda_{k,t}-1\right)}{\Gamma\left(\lambda_{k,t}-1\right)\Gamma\left(Y_{k,t}\right)}\upsilon_k^{\lambda_{k,t}}\left(1-\upsilon_k\right)^{Y_{k,t}}\right)\mathbf{1}\left(S_t=k\right)$$
$$\Theta = \left(\theta_1,...,\theta_{m+1}\right),$$
$$\text{with } \theta_k = \left(\alpha_{k,0},\alpha_{k,1},\beta_k,\left(\gamma_{k,1},...,\gamma_{k,q}\right),\upsilon_k\right)' \text{ for } k=1,...,m+1$$

For the second case of Generalized Poisson *INGARCHX* model with $Y_{S_t,t}|\mathscr{F}_{t-1} \rightsquigarrow \mathscr{G}\mathscr{P}\left((1-\phi_{S_t})\,\lambda_{S_t,t}, \phi_{S_t}\right)$, the conditional likelihood function of the n observations $Z_1,...,Z_n$ conditionally on the presample values, is given by

$$L(\Theta) = \prod_{t=1}^{n}\sum_{k=1}^{m+1}\left(\frac{(1-\phi_k)\,\lambda_{k,t}\left[(1-\phi_k)\,\lambda_{k,t}+\phi_k Y_{k,t}\right]^{Y_{k,t}-1}e^{-\left[(1-\phi_k)\lambda_{k,t}+\phi_k Y_{k,t}\right]}}{Y_{k,t}\,!}\right)\mathbf{1}\left(S_t=k\right)$$
$$\Theta = \left(\theta_1,...,\theta_{m+1}\right),$$
$$\text{with } \theta_k = \left(\alpha_{k,0},\alpha_{k,1},\beta_k,\left(\gamma_{k,1},...,\gamma_{k,q}\right),\phi_k\right)' \text{ for } k=1,...,m+1$$

When the vector of break or change points $(c_1,c_2,...,c_m)'$ is unknown, the *CLS* and *CML* methods mentioned above cannot be applied immediately since the change points parameter lies in an indicator function. Therefore, we need to estimate $(c_1,c_2,...,c_m)'$ first. A commonly used approach to estimate the multiple breakpoints is the so-called search minimum least squares (*SMLS*) algorithm proposed by Perron *et al* (1998). Motivated by the work of Perron *et al* (1998), we proposed an adapted version of Perron's algorithm based on the (*SMLS*) algorithm to estimate the $(m+1)$ change points in our formulation.

4. REFERENCES

[1] Fokianos, K. *Count time series models*. In Handbook of Statistics-Time Series Analysis : Methods and Applications, Amsterdam : Elsevier, 315-347, 2012.

[2] Chen, C. WS. and Khamthong. K. Bayesian modeling of nonlinear negative binomial integer-valued GARCHX models. *Statistical Modeling*, Numéro :20(6) 537–561, 2019.

[3] Lee, Sangyeol. and Lee, Sangjo. Change Point Test for the Conditional Mean of Time Series of Counts Based on Support Vector Regression. *Entropy*, 2021.

[4] Mao, H., Zhu, Fukung. and Cui. Y. A generalized mixture integer-valued GARCH model. *Statistical Methods & Applications*, Numéro :29 527–552, 2019.

[5] Perron, p. and Bai. J. Estimatin and testing linear models with multiple structural changes. *Econoometrica*, Numéro :1 47–78, 1998.

[6] Wang, C. Liu, H. Yao, J. F. Davis, R.A. and Li. W.K. Self-Excited Threshold Poisson Autoregression. *Journal of the American Statistical Association*, Numéro :109(506) 777–787, 2014.