C. R. Acad. Sci. Paris, Ser. I 355 (2017) 99–104

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Probabilités/Statistique

Estimateur du quasi-maximum de vraisemblance géométrique d'une classe générale de modèles de séries chronologiques à valeurs entières

Geometric quasi-maximum likelihood estimation for a general class of integer-valued time series models

Abdelhakim Aknouche, Sara Bendjeddou

Faculté de mathématiques, Université des sciences et de la technologie Houari-Boumédiène (USTHB), Algérie

I N F O   A R T I C L E

Historique de l'article :
Reçu le 4 juillet 2016
Accepté après révision le 24 novembre 2016
Disponible sur Internet le 8 décembre 2016

Présenté par le comité de rédaction

R É S U M É

Cette note établit la convergence presque sûre et la normalité asymptotique de l'estimateur du quasi-maximum de vraisemblance géométrique d'une classe générale de modèles de séries chronologiques à valeurs entières. Dans cette classe, le modèle spécifie seulement la moyenne conditionnelle du processus, sous une forme paramétrique générale. Une comparaison avec l'estimateur du quasi-maximum de vraisemblance poissonnien, en termes d'efficacité asymptotique relative, est considérée.

A B S T R A C T

This note establishes the consistency and the asymptotic normality of the geometric quasi-maximum-likelihood estimate ($QMLE$) of a general class of integer-valued time series models. In this class, only the conditional mean is specified in a general parametric form. Comparison with the Poisson $QMLE$ on some particular models, with regard to asymptotic relative efficiency, is considered.

1. Introduction

Un intérêt considérable a été porté au cours de ces trois dernières décennies aux modèles de séries chronologiques à valeurs entières (voir, par exemple, [5]). De nombreux modèles ont été introduits pour tenir compte des spécificités des séries chronologiques à valeurs entières, que sont notamment la surdispersion, la surfréquence de zéros, l'asymétrie et une structure d'autocorrélation positive. Deux classes générales de modèles ont attiré l'attention : i) les modèles basés sur

Adresses e-mail : aknouche_ab@yahoo.com (A. Aknouche), bendjeddousara@yahoo.com (S. Bendjeddou).

http://dx.doi.org/10.1016/j.crma.2016.11.006
1631-073X/

la régression discrète (comme, par exemple, l'autorégression de Poisson, ou l'autorégression en binomiale négative), dont l'exemple le plus représentatif est le modèle $GARCH$ entier (ex. [8], [6], [11], [4]), et ii) les modèles basés sur les équations aux différences stochastiques faisant intervenir l'opérateur d'amincissement (*thinning operator*) dont l'exemple typique est le modèle $INAR$ (INteger AR, [2]). Ahmad et Francq [1] ont considéré une classe de modèles plus générale, englobant les deux classes citées précédemment, dont la moyenne conditionnelle est fonction paramétrée du passé infini du processus sous-jacent. Pour cette classe, dont seule la moyenne conditionnelle est à spécifier, Ahmad et Francq [1] ont établi la convergence presque sûre et la normalité asymptotique de l'estimateur du quasi-maximum de vraisemblance poissonnien ($PQMLE$). Cet estimateur est calculé comme si la distribution conditionnelle des observations était poissonnienne. Le $PQMLE$ présente plusieurs qualités statistiques désirables. D'abord, il est asymptotiquement efficace lorsque la loi conditionnelle des observations est poissonnienne. Ensuite, lorsque la moyenne et la variance conditionnelles du modèle sont proportionnelles, il est asymptotiquement efficace dans la classe de tous les estimateurs du quasi-maximum de vraisemblance (QML) appartenant à la famille exponentielle linéaire (cf. [7]). Enfin, il est robuste devant une mauvaise spécification de la loi conditionnelle du modèle. Cependant, en dépit de ces avantages et de l'importance de la loi de Poisson pour les phénomènes discrets, cette loi, étant équidispersée, s'ajuste mal aux séries réelles surdispersées que l'on observe très fréquemment en pratique. Ainsi, le $PQMLE$ n'est pas asymptotiquement efficace si l'on l'applique à des phénomènes à valeurs entières affectés de surdispersion. Aussi, le recours à un estimateur du QML calculé sur la base d'une distribution surdispersée et appartenant à la famille exponentielle pourrait être un complément intéressant au $PQMLE$.

Pour la classe de modèles étudiés par Ahmad et Francq [1], nous proposons dans ce travail un estimateur du QML géométrique ($GQMLE$), calculé comme si la distribution conditionnelle du processus observable était géométrique. Nous en établissons la convergence presque sûre et la normalité asymptotique sous un jeu d'hypothèses assez faibles, en particulier sans la spécification entière de la distribution conditionnelle du modèle. Bien entendu, cet estimateur est asymptotiquement efficace lorsque la distribution conditionnelle des observations est géométrique. De plus, puisqu'il est fondé sur une loi surdispersée, l'estimateur du $GQML$ peut avoir un meilleur niveau d'efficacité asymptotique relative que le $PQMLE$ en présence de séries réelles surdispersées.

Le reste de cette note est structuré comme suit. La Section 2 définit le modèle sous-jacent et le $GQMLE$ correspondant, puis établit la consistance et la normalité asymptotique de ce dernier. La Section 3 compare le $GQMLE$ et le $PQMLE$ sur des hypothèses et des modèles spécifiques.

2. Estimateur du QML géométrique pour des modèles à valeurs entières

Soit un paramètre inconnu $\theta_0 \in \Theta \subset \mathbb{R}^m$ ($m \in \mathbb{N}^* = \mathbb{N}/\{0\}$) et une fonction mesurable à valeurs réelles $\lambda : \mathbb{N}^\infty \times \Theta \to]0, \infty[$. Considérons un processus aléatoire à valeurs entières, $\{X_t, t \in \mathbb{Z}\}$, défini sur un espace probabilisé $(\Omega, \mathcal{F}, P)$, avec une moyenne conditionnelle de la forme

$$E\left(X_t/\mathcal{F}_{t-1}\right) = \lambda\left(X_{t-1}, X_{t-2}, \ldots; \theta_0\right) := \lambda_t\left(\theta_0\right) := \lambda_t, \quad t \in \mathbb{Z}, \tag{2.1}$$

où $\mathcal{F}_t$ est la tribu générée par $\{X_t, X_{t-1}, \ldots\}$. La généralité du modèle (2.1) provient de la forme quelconque de la fonction λ, et aussi du fait que seule la moyenne conditionnelle est à spécifier. Il englobe plusieurs classes importantes de modèles de séries chronologiques à valeurs entières, notamment le modèle $GARCH$ poissonnien stable [6,8], le modèle $GARCH$ en binomiale négative stable [11,4] et le modèle AR entier ($INAR$, [2]). Dans le reste de cette note, nous nous intéressons à l'estimation de θ_0 pour une série $X_1, X_2, \ldots, X_n$ satisfaisant (2.1) avec l'hypothèse suivante :

H0 *Le processus $\{X_t, t \in \mathbb{Z}\}$ spécifié par (2.1) est strictement stationnaire et ergodique.*

Pour certaines classes particulières de processus satisfaisant (2.1), l'hypothèse **H0** s'exprime plus explicitement par une condition de stabilité sur θ_0. Pour tout paramètre générique $\theta \in \Theta$, la fonction moyenne conditionnelle définie par $\lambda_t(\theta) := \lambda\left(X_{t-1}, X_{t-2}, \ldots; \theta\right)$ coïncide avec la moyenne conditionnelle donnée par (2.1) lorsque $\theta = \theta_0$. Étant donné des valeurs initiales arbitrairement fixées $\widetilde{X}_0, \widetilde{X}_{-1}, \ldots$, soit $\widetilde{\lambda}_t(\theta) = \lambda\left(X_{t-1}, X_{t-2}, \ldots X_1, \widetilde{X}_0, \widetilde{X}_{-1}, \ldots; \theta\right)$ ($t \in \mathbb{N}^*$), une approximation de $\lambda_t(\theta)$ servant à calculer effectivement un estimateur de θ_0. En se basant sur la loi géométrique de paramètre $\frac{1}{1+\widetilde{\lambda}_t(\theta)}$, la fonction de vraisemblance géométrique du modèle (2.1) est donnée, après transformation logarithmique, par

$$\widetilde{L}_{G,n}(\theta) = \frac{1}{n}\sum_{t=1}^{n} \widetilde{l}_t(\theta) \text{ avec } \widetilde{l}_t(\theta) = \log\left(\frac{1}{1+\widetilde{\lambda}_t(\theta)}\right) + X_t \log\left(\frac{\widetilde{\lambda}_t(\theta)}{1+\widetilde{\lambda}_t(\theta)}\right), \quad \theta \in \Theta. \tag{2.2}$$

Un estimateur du quasi-maximum de vraisemblance géométrique ($GQMLE$) de θ_0 est alors le maximum de $\widetilde{L}_{G,n}(\theta)$ sur Θ, c.-à-d. une solution mesurable du problème $\widehat{\theta}_G = \arg\max_{\theta\in\Theta}\left(\widetilde{L}_{G,n}(\theta)\right)$ pour un certain espace paramétrique Θ. Par opposition, l'estimateur du QML Poissonnien ($PQMLE$, cf. [1]) est solution du problème $\widehat{\theta}_P = \arg\max_{\theta\in\Theta}\left(\widetilde{L}_{P,n}(\theta)\right)$ où $\widetilde{L}_{P,n}(\theta) = \frac{1}{n}\sum_{t=1}^{n}\left(-\widetilde{\lambda}_t(\theta) + X_t \log\left(\widetilde{\lambda}_t(\theta)\right)\right)$. Le choix des valeurs initiales $\widetilde{X}_0, \widetilde{X}_{-1}, \ldots$ est sans effet asymptotiquement, mais peut être important en échantillons finis (cf. [1]). Par ailleurs, la fonction de vraisemblance donnée par (2.2) appartient à la famille exponentielle. Elle coïncide avec la fonction de quasi-vraisemblance de Wedderburn [9] fondée sur l'hypothèse du modèle linéaire généralisé (Generalized Linear Model, GLM)

$$\operatorname{Var}\left(X_t/\mathcal{F}_{t-1}\right) = E\left(X_t/\mathcal{F}_{t-1}\right)\left(1 + E\left(X_t/\mathcal{F}_{t-1}\right)\right) = \lambda_t\left(1+\lambda_t\right). \tag{2.3}$$

Pour étudier la convergence presque sûre du $GQMLE$, considérons les hypothèses suivantes.

H1 *La fonction* $\theta \mapsto \lambda_t(\theta)$ *est continue* $p.s.$ *et pour tout* $t \in \mathbb{Z}$: $\lambda_t(\theta) > c$ *et* $\widetilde{\lambda}_t(\theta) > c$, $p.s.$, *où* $c > 0$.
H2 $a_t \underset{t\to\infty}{\overset{p.s.}{\to}} 0$ *et* $a_t X_t \underset{t\to\infty}{\overset{p.s.}{\to}} 0$ *où* $a_t = \sup_{\theta\in\Theta}\left|\widetilde{\lambda}_t(\theta) - \lambda_t(\theta)\right|$.
H3 $E\left(X_t^{\delta}\right) < \infty$ pour un certain $\delta > 1$.
H4 $\lambda_t(\theta) = \lambda_t(\theta_0)$ $p.s.$ *si et seulement si* $\theta = \theta_0$.
H5 Θ *est compact.*

Les hypothèses **H1–H5** sont standard. Certaines peuvent être explicitées pour des classes particulières de modèles de la forme (2.1). Elles sont similaires à celles faites par Ahmad et Francq [1] pour la convergence forte de leur $PQMLE$.

Théorème 2.1. *Sous* (2.1) *et* ***H0–H5***, $\widehat{\theta}_G \underset{n\to\infty}{\overset{p.s.}{\to}} \theta_0$.

Le résultat précédant montre que, comme le $PQMLE$, le $GQMLE$ est robuste devant une mauvaise spécification de la loi conditionnelle du modèle, pourvu que la moyenne conditionnelle soit bien spécifiée. Ceci n'est pas surprenant, puisque la fonction de vraisemblance géométrique appartient à la famille exponentielle (cf. [10], [7]). Soit $l_t(\theta)$ et $L_{G,n}(\theta)$ les variables aléatoires obtenues à partir de $\widetilde{l}_t(\theta)$ et $\widetilde{L}_{G,n}(\theta)$, respectivement, en remplaçant $\widetilde{\lambda}_t(\theta)$ par $\lambda_t(\theta)$. Pour établir la normalité asymptotique du $GQMLE$, considérons les hypothèses de régularité supplémentaires suivantes.

H6 *Les quantités* c_t, $c_t X_t$, $a_t d_t$, $a_t d_t X_t$ *et* $b_t d_t X_t$ *sont d'ordre* $O\left(t^{-\tau}\right)$ $p.s.$ *pour un certain* $\tau > 1/2$ *où* $b_t = \sup_{\theta\in\Theta}\left|\widetilde{\lambda}_t^2(\theta) - \lambda_t^2(\theta)\right|$, $c_t = \sup_{\theta\in\Theta}\left\|\frac{\partial\left(\widetilde{\lambda}_t(\theta)-\lambda_t(\theta)\right)}{\partial\theta}\right\|$ *et* $d_t = \sup_{\theta\in\Theta}\max\left(\left\|\frac{1}{\widetilde{\lambda}_t(\theta)(1+\widetilde{\lambda}_t(\theta))}\frac{\partial\widetilde{\lambda}_t(\theta)}{\partial\theta}\right\|, \left\|\frac{1}{\lambda_t(\theta)(1+\lambda_t(\theta))}\frac{\partial\lambda_t(\theta)}{\partial\theta}\right\|\right)$.
H7 *La vraie valeur* θ_0 *est à l'intérieur de* Θ.
H8 *La variance conditionnelle* $v_t(\theta_0) := Var\left(X_t/\mathcal{F}_{t-1}\right) = E\left(X_t^2/\mathcal{F}_{t-1}\right) - \lambda_t^2(\theta_0)$ *est finie* $p.s.$
H9 Les dérivées $\frac{\partial^2\lambda_t(\theta)}{\partial\theta\partial\theta'}$ et $\frac{\partial^2\widetilde{\lambda}_t(\theta)}{\partial\theta\partial\theta'}$ existent et sont continues $p.s.$, *les matrices* $I_G = E\left(\frac{v_t(\theta_0)}{\lambda_t^2(\theta_0)(1+\lambda_t(\theta_0))^2}\frac{\partial\lambda_t(\theta_0)}{\partial\theta}\frac{\partial\lambda_t(\theta_0)}{\partial\theta'}\right)$ *et* $J_G = E\left(\frac{1}{\lambda_t(\theta_0)(1+\lambda_t(\theta_0))}\frac{\partial\lambda_t(\theta_0)}{\partial\theta}\frac{\partial\lambda_t(\theta_0)}{\partial\theta'}\right)$ *sont finies et* J_G *est inversible.*
H10 *Il existe un voisinage* $V(\theta_0)$ *de* θ_0 *tel que* $E\left(\sup_{\theta\in V(\theta_0)}\left\|\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'}\right\|\right) < \infty$.

Théorème 2.2. *Sous* (2.1) *et* ***H0–H10***,

$$\sqrt{n}\left(\widehat{\theta}_G - \theta_0\right) \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} N\left(0, J_G^{-1} I_G J_G^{-1}\right). \tag{2.4}$$

Comme corollaire au Théorème 2.2, un estimateur consistant de la variance asymptotique $J_G^{-1} I_G J_G^{-1}$ est $\widehat{J}_G^{-1}\widehat{I}_G\widehat{J}_G^{-1}$ avec $\widehat{I}_G = \frac{1}{n}\sum_{t=1}^n\left(\frac{X_t-\widetilde{\lambda}_t(\widehat{\theta}_G)}{\widetilde{\lambda}_t(\widehat{\theta}_G)(1+\widetilde{\lambda}_t(\widehat{\theta}_G))}\right)^2\frac{\partial\widetilde{\lambda}_t(\widehat{\theta}_G)}{\partial\theta}\frac{\partial\widetilde{\lambda}_t(\widehat{\theta}_G)}{\partial\theta'}$ et $\widehat{J}_G = \frac{1}{n}\sum_{t=1}^n\frac{1}{\widetilde{\lambda}_t(\widehat{\theta}_G)(1+\widetilde{\lambda}_t(\widehat{\theta}_G))}\frac{\partial\widetilde{\lambda}_t(\widehat{\theta}_G)}{\partial\theta}\frac{\partial\widetilde{\lambda}_t(\widehat{\theta}_G)}{\partial\theta'}$.

3. Comparaison avec l'estimateur du QML poissonnien

Lorsque la loi conditionnelle de $\{X_t, t\in\mathbb{Z}\}$ est géométrique de paramètre $\frac{1}{1+\lambda_t}$ (i.e. $X_t/\mathcal{F}_{t-1} \sim \mathcal{G}\left(\frac{1}{1+\lambda_t}\right)$), l'estimateur $\widehat{\theta}_G$ n'est autre que l'estimateur du maximum de vraisemblance. Étant alors asymptotiquement efficace, il ne peut être asymptotiquement moins précis que l'estimateur $\widehat{\theta}_P$. Dans ce cas, $J_G = I_G$ et (2.4) devient $\sqrt{n}\left(\widehat{\theta}_G - \theta_0\right) \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} N\left(0, J_G^{-1}\right)$. Une autre propriété importante est que, sous la condition GLM suivante,

$$\mathrm{Var}\left(X_t/\mathcal{F}_{t-1}\right) = \rho^2 E\left(X_t/\mathcal{F}_{t-1}\right)\left(1 + E\left(X_t/\mathcal{F}_{t-1}\right)\right) = \rho^2\lambda_t\left(1+\lambda_t\right) \text{ pour un certain } \rho^2 > 0, \tag{3.1}$$

qui généralise (2.3), l'estimateur $\widehat{\theta}_G$ est asymptotiquement efficace dans la classe de tous les estimateurs du QML appartenant à la famille exponentielle linéaire (cf. [7]). Du fait que la vraisemblance poissonnienne appartient à cette famille, il s'ensuit que, sous (3.1), $\widehat{\theta}_G$ est plus efficace asymptotiquement que le $PQMLE$. Dans ce cas, on a $\sqrt{n}\left(\widehat{\theta}_G - \theta_0\right) \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} N\left(0, \rho^2 J_G^{-1}\right)$. Par dualité, si la loi conditionnelle de $\{X_t, t\in\mathbb{Z}\}$ est poissonnienne de paramètre λ_t, alors (cf. [1]) $\widehat{\theta}_P$ est asymptotiquement efficace, et en particulier plus efficace asymptotiquement que $\widehat{\theta}_G$. De plus, sous la condition GLM poissonnienne

$$\mathrm{Var}\left(X_t/\mathcal{F}_{t-1}\right) = \rho^2 E\left(X_t/\mathcal{F}_{t-1}\right) \text{ pour un certain } \rho^2 > 0, \tag{3.2}$$

qui se vérifie pour la loi de Poisson avec $\rho^2 = 1$, l'estimateur $\widehat{\theta}_P$ est asymptotiquement efficace dans la classe de tous les estimateurs du QML appartenant à la famille exponentielle linéaire, et ainsi il est plus efficace asymptotiquement que $\widehat{\theta}_G$. Un exemple de GLM spécifique est obtenu avec le modèle $GARCH$ entier stable défini par $X_t/\mathcal{F}_{t-1} \sim \text{Loi}(\lambda_t)$ et $\lambda_t := \lambda_t(\theta_0) = \omega_0 + \sum_{i=1}^{q} \alpha_{0i} X_{t-i} + \sum_{j=1}^{p} \beta_{0j} \lambda_{t-j}(\theta_0)$ où $\omega_0 > 0$, $\alpha_{0i} \geq 0$, $\beta_{0j} \geq 0$ et $\sum_{j=1}^{p} \alpha_{0i} + \beta_{0j} < 1$. Loi (λ_t) désigne la loi de Poisson $\mathcal{P}(\lambda_t)$, la loi géométrique $\mathcal{G}\left(\frac{1}{1+\lambda_t}\right)$ ou toute autre loi discrète pourvu que (2.1) soit vérifiée. Un autre exemple de la forme (2.1), mais pour lequel aucun des estimateurs $\widehat{\theta}_G$ et $\widehat{\theta}_P$ n'est en général plus efficace asymptotiquement que l'autre, est le modèle $INAR(1)$. Ce processus est défini par $X_t = \sum_{i=1}^{X_{t-1}} Y_i + \varepsilon_t$, où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires entières indépendantes et identiquement distribuées (iid) de moyenne $\omega_0 > 0$ et de variance $\sigma_0^2 > 0$ et où $\{Y_i, i \in \mathbb{N}\}$ est une suite iid de loi de Bernoulli de paramètre $\alpha_0 \in]0, 1[$. En effet, quelle que soit la loi marginale de l'innovation $\{\varepsilon_t, t \in \mathbb{Z}\}$, qu'elle soit géométrique, poissonnienne, ou toute autre loi discrète, le modèle $INAR(1)$ vérifie une relation GLM affine donnée par : $\text{Var}(X_t/\mathcal{F}_{t-1}) = (1-\alpha_0) E(X_t/\mathcal{F}_{t-1}) + \sigma_0^2 - (1-\alpha_0)\omega_0$. Cette relation ne peut se réduire à aucune des relations GLM (3.1) (géométrique) et (3.2) (poissonnienne), à moins que $\frac{\sigma_0^2}{\omega_0} = 1 - \alpha_0$, auquel cas ε_t devrait être sous-dispersée et $\widehat{\theta}_P$ serait plus efficace asymptotiquement que $\widehat{\theta}_G$.

Remerciements

Nous sommes très reconnaissants à un rapporteur anonyme pour son examen minutieux et ses remarques et suggestions constructives. Nous sommes également profondément redevables à Christian Francq pour son aide substantielle dans la révision du manuscrit, mais nous sommes seuls responsables d'éventuelles erreurs.

Annexe A. Preuves des résultats

Preuve du Théorème 2.1. Suivant l'approche de Wald, la preuve du Théorème 2.1 repose sur les deux lemmes suivants.

Lemme A.1. *Sous* ***H1–H2***, $\sup_{\theta\in\Theta} \left| L_{G,n}(\theta) - \widetilde{L}_{G,n}(\theta) \right| \underset{n\to\infty}{\overset{p.s.}{\to}} 0$.

Preuve. En utilisant l'inégalité $\log(x) \leq x - 1$, le fait que $\widetilde{\lambda}_t(\theta) > 0$, les hypothèses **H1–H2**, puis le lemme de Césaro, il s'ensuit que

$$\begin{aligned}\sup_{\theta\in\Theta} \left| L_{G,n}(\theta) - \widetilde{L}_{G,n}(\theta) \right| &= \frac{1}{n} \sup_{\theta\in\Theta} \left| \sum_{t=1}^{n} \left(\log\left(\frac{\lambda_t(\theta) - \widetilde{\lambda}_t(\theta)}{1+\widetilde{\lambda}_t(\theta)} + 1 \right) + X_t \log\left(\frac{\widetilde{\lambda}_t(\theta) - \lambda_t(\theta)}{\lambda_t(\theta)(1+\widetilde{\lambda}_t(\theta))} + 1 \right) \right) \right| \\ &\leq \frac{1}{n} \sum_{t=1}^{n} \left(a_t + X_t a_t \frac{1}{c} \right) \underset{n\to\infty}{\overset{p.s.}{\to}} 0. \qquad \square\end{aligned}$$

Lemme A.2. *Sous* ***H0–H1*** *et* ***H3–H4*** *: i)* $E(l_1(\theta_0)) < \infty$, *ii)* $E(l_1(\theta_0)) \geq E(l_1(\theta))$ *pour tout* $\theta \in \Theta$, *et iii)* $E(l_1(\theta)) = E(l_1(\theta_0)) \Rightarrow \theta = \theta_0$.

Preuve. Sous **H1**, les variables $\log\left(\frac{1}{1+\lambda_t(\theta)}\right)$ et $\log\left(\frac{\lambda_t(\theta)}{1+\lambda_t(\theta)}\right)$ étant bornées, elles admettent des moments finis à tout ordre. Par les inégalités de Jensen et de Hölder, ainsi que **H3**, il en découle que

$$\begin{aligned}|E(l_1(\theta_0))| < E(|l_1(\theta_0)|) &\leq E\left(\left| \log\left(\frac{1}{1+\lambda_t(\theta_0)} \right) \right| \right) + E\left(\left| X_t \log\left(\frac{\lambda_t(\theta_0)}{1+\lambda_t(\theta_0)} \right) \right| \right) \\ &\leq E\left(\left| \log\left(\frac{1}{1+\lambda_t(\theta_0)} \right) \right| \right) + \left(E\left(X_t^{\delta}\right) \right)^{1/\delta} \left(E\left(\left| \log \frac{\lambda_t(\theta_0)}{1+\lambda_t(\theta_0)} \right|^{\frac{\delta}{\delta-1}} \right) \right)^{\frac{\delta-1}{\delta}} < \infty. \end{aligned} \tag{A.1}$$

En utilisant de nouveau le fait que $\log(x) \leq x - 1$, nous avons l'inégalité suivante

$$\begin{aligned}E(l_1(\theta) - l_1(\theta_0)) &= E\left(\log\left(\frac{1+\lambda_t(\theta_0)}{1+\lambda_t(\theta)} \right) + X_t \log\left(\frac{\lambda_t(\theta)(1+\lambda_t(\theta_0))}{\lambda_t(\theta_0)(1+\lambda_t(\theta))} \right) \right) \\ &\leq E\left(\left(\frac{1+\lambda_t(\theta_0)}{1+\lambda_t(\theta)} - 1 \right) + X_t \left(\frac{\lambda_t(\theta)(1+\lambda_t(\theta_0))}{\lambda_t(\theta_0)(1+\lambda_t(\theta))} - 1 \right) \right) = 0, \end{aligned} \tag{A.2}$$

où l'on peut observer par (A.1) et (A.2) que $E(l_1(\theta) - l_1(\theta_0)) \in [-\infty, 0]$ et alors $E(l_1(\theta)) < E(l_1(\theta_0))$ pour tout $\theta \neq \theta_0$. Enfin, l'inégalité (A.2) se réduit à une égalité si et seulement si $E\left(\log\left(\frac{1+\lambda_t(\theta_0)}{1+\lambda_t(\theta)} \right) + X_t \log\left(\frac{\lambda_t(\theta)(1+\lambda_t(\theta_0))}{\lambda_t(\theta_0)(1+\lambda_t(\theta))} \right) \right) = 0$, ce qui se vérifie si et seulement si $\lambda_t(\theta) = \lambda_t(\theta_0)$ et alors par l'hypothèse d'identifiabilité **H4**, si et seulement si $\theta = \theta_0$. $\square$

Pour compléter la preuve du théorème, soit $V_k(\overline{\theta})$ ($k \in \mathbb{N}^*$, $\overline{\theta} \in \Theta$) la boule ouverte de centre $\overline{\theta}$ et de rayon $1/k$. Puisque $\sup_{\theta \in V_k(\overline{\theta}) \cap \Theta} l_t(\theta)$ est fonction mesurable des termes de $\{X_t, t \in \mathbb{Z}\}$, processus strictement stationnaire et ergodique sous **H0**, il s'ensuit que $\left\{\sup_{\theta \in V_k(\overline{\theta}) \cap \Theta} l_t(\theta), t \in \mathbb{Z}\right\}$ est également strictement stationnaire et ergodique, où l'on peut constater par le Lemme A.2 que $E\left(\sup_{\theta \in V_k(\overline{\theta}) \cap \Theta} l_t(\theta)\right) \in [-\infty, +\infty[$. Ainsi, au vu du Lemme A.1 et du théorème ergodique (ex. [3]), il en résulte que

$$\limsup_{n\to\infty} \sup_{\theta \in V_k(\overline{\theta}) \cap \Theta} \widetilde{L}_{G,n}(\theta) = \limsup_{n\to\infty} \sup_{\theta \in V_k(\overline{\theta}) \cap \Theta} L_{G,n}(\theta) \le E\left(\sup_{\theta \in V_k(\overline{\theta}) \cap \Theta} l_1(\theta)\right).$$

Par le théorème de Beppo–Levi, $E\left(\sup_{\theta \in V_k(\overline{\theta}) \cap \Theta} l_1(\theta)\right)$ décroît vers $E\left(l_1\left(\overline{\theta}\right)\right)$ lorsque $k \to \infty$. Donc, nous avons montré que, d'après le Lemme A.1 et le Lemme A.2, il existe, pour tout $\overline{\theta} \neq \theta_0$, un voisinage $V\left(\overline{\theta}\right)$ tel que

$$\limsup_{n\to\infty} \sup_{\theta \in V_k(\overline{\theta}) \cap \Theta} \widetilde{L}_{G,n}(\theta) < \limsup_{n\to\infty} \widetilde{L}_{G,n}(\theta_0) = \limsup_{n\to\infty} L_{G,n}(\theta_0) = E\left(l_1(\theta_0)\right).$$

Ainsi, par des arguments standards, la preuve du théorème s'achève en exploitant l'hypothèse **H5** de compacité de Θ. □

Preuve du Théorème 2.2. D'abord, par **H6**, nous avons

$$\sqrt{n} \sup_{\theta \in \Theta} \left\| \frac{\partial \widetilde{L}_{G,n}(\theta)}{\partial \theta} - \frac{\partial L_{G,n}(\theta)}{\partial \theta} \right\| \le \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(c_t + a_t d_t + X_t \left(\frac{c_t}{c} + \frac{(a_t + b_t) d_t}{c^2} \right) \right) = o_{p.s.}(1).$$

Ensuite, comme par **H7** et le Théorème 2.1, $\widehat{\theta}_G$ ne peut être sur le bord de Θ pour n assez grand, il s'ensuit par la formule de Taylor que, pour un certain $\theta^* \in \Theta$ compris entre $\widehat{\theta}_G$ et θ_0,

$$0 = \sqrt{n} \frac{\partial \widetilde{L}_{G,n}(\widehat{\theta}_G)}{\partial \theta} = \sqrt{n} \frac{\partial L_{G,n}(\widehat{\theta}_G)}{\partial \theta} + o_{p.s.}(1) = \sqrt{n} \frac{\partial L_{G,n}(\theta_0)}{\partial \theta} + \sqrt{n} \frac{\partial^2 L_{G,n}(\theta^*)}{\partial \theta \partial \theta'} \left(\widehat{\theta}_G - \theta_0\right) + o_{p.s.}(1). \tag{A.3}$$

Enfin, en se basant sur (A.3), la preuve du théorème s'établit en démontrant les deux lemmes suivants.

Lemme A.3. *Sous **H8**–**H9**,* $\sqrt{n} \frac{\partial L_{G,n}(\theta_0)}{\partial \theta} \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} N(0, I_G)$.

Preuve. Il est clair que $\left\{\frac{\partial l_t(\theta_0)}{\partial \theta}, t \in \mathbb{Z}\right\}$ est une différence de martingale par rapport à $\{\mathcal{F}_t, t \in \mathbb{Z}\}$ où $\frac{\partial l_t(\theta_0)}{\partial \theta} = \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{X_t - \lambda_t(\theta_0)}{\lambda_t(\theta_0)(1+\lambda_t(\theta_0))}$ et $\sqrt{n} \frac{\partial L_{G,n}(\theta_0)}{\partial \theta} = \sum_{t=1}^{n} \frac{1}{\sqrt{n}} \frac{\partial l_t(\theta_0)}{\partial \theta}$. Par **H8**–**H9**, $E\left(\frac{\partial l_t(\theta_0)}{\partial \theta} \frac{\partial l_t(\theta_0)}{\partial \theta'}\right) = I_G$, et ainsi le lemme s'ensuit par le théorème central limite des martingales (ex. [3]). □

Lemme A.4. *Sous **H8**–**H10**,* $\frac{\partial^2 L_{G,n}(\theta^*)}{\partial \theta \partial \theta'} \underset{n\to\infty}{\overset{p.s.}{\to}} J_G$.

Preuve. Soit $V_k(\theta_0)$ ($k \in \mathbb{N}^*$) la boule ouverte de centre θ_0 et de rayon $1/k$, où k est supposé assez grand pour que $V_k(\theta_0)$ soit contenu dans $V(\theta_0)$, qui est défini par **H10**. Supposons également que n soit assez grand pour que θ^* soit dans $V_k(\theta_0)$. Donc, par stationnarité et ergodicité de $\left\{\sup_{\theta \in V_k(\theta_0)} \left| \frac{\partial^2 l_t(\theta)}{\partial \theta_i \partial \theta_j} - E\left(\frac{\partial^2 l_t(\theta_0)}{\partial \theta_i \partial \theta_j}\right) \right|, t \in \mathbb{Z}\right\}$, nous avons

$$\left| \frac{\partial^2 L_{G,n}(\theta^*)}{\partial \theta \partial \theta'} - J_G \right| \le \frac{1}{n} \sum_{t=1}^{n} \sup_{\theta \in V_k(\theta_0)} \left| \frac{\partial^2 l_t(\theta)}{\partial \theta_i \partial \theta_j} - E\left(\frac{\partial^2 l_t(\theta_0)}{\partial \theta_i \partial \theta_j}\right) \right| \underset{n\to\infty}{\overset{p.s.}{\to}} E\left(\sup_{\theta \in V_k(\theta_0)} \left| \frac{\partial^2 l_t(\theta)}{\partial \theta_i \partial \theta_j} - E\left(\frac{\partial^2 l_t(\theta_0)}{\partial \theta_i \partial \theta_j}\right) \right|\right).$$

Au vu de **H10**, le théorème de convergence dominée de Lebesgue entraîne que

$$\lim_{k\to\infty} E\left(\sup_{\theta \in V_k(\theta_0)} \left| \frac{\partial^2 l_t(\theta)}{\partial \theta_i \partial \theta_j} - E\left(\frac{\partial^2 l_t(\theta_0)}{\partial \theta_i \partial \theta_j}\right) \right|\right) = E\left(\lim_{k\to\infty} \sup_{\theta \in V_k(\theta_0)} \left| \frac{\partial^2 l_t(\theta)}{\partial \theta_i \partial \theta_j} - E\left(\frac{\partial^2 l_t(\theta_0)}{\partial \theta_i \partial \theta_j}\right) \right|\right) = 0,$$

ce qui complète la preuve du lemme et par là même celle du théorème. □

Références

[1] A. Ahmad, C. Francq, Poisson QMLE of count time series models, J. Time Ser. Anal. 37 (2016) 291–314.
[2] M.A. Al-Osh, A.A. Alzaid, First-order integer-valued autoregressive ($INAR(1)$) process, J. Time Ser. Anal. 8 (1987) 261–275.
[3] P. Billingsley, Probability and Measure, 3rd edition, John Wiley & Sons, 2008.
[4] V. Christou, K. Fokianos, Quasi-likelihood inference for negative binomial time series models, J. Time Ser. Anal. 35 (2014) 55–78.

[5] R.A. Davis, S.H. Holan, R. Lund, N. Ravishanker, Handbook of Discrete-Valued Time Series, Chapman and Hall, 2016.
[6] R. Ferland, A. Latour, D. Oraichi, Integer-valued $GARCH$ process, J. Time Ser. Anal. 27 (2006) 923–942.
[7] C. Gourieroux, A. Monfort, A. Trognon, Pseudo maximum likelihood methods: theory, Econometrica 52 (1984) 681–700.
[8] A. Heinen, Modelling time series count data: an autoregressive conditional Poisson model, CORE Discussion Paper 2003/62, Université catholique de Louvain, 2003.
[9] R.W. Wedderburn, Quasi-likelihood functions, generalized linear models, and the Gauss–Newton method, Biometrika 61 (1974) 439–447.
[10] H. White, Maximum likelihood of misspecified models, Econometrica 50 (1982) 1–25.
[11] F. Zhu, A negative binomial integer-valued $GARCH$ model, J. Time Ser. Anal. 32 (2011) 54–67.