

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediene
Faculté de Mathématiques



Thèse

présentée pour l'obtention du Grade de Docteur en Sciences

en Mathématiques

Spécialité : **Recherche Opérationnelle : Méthodes Stochastiques**

Présentée par : BENDJEDDOU Sara

Thème

**Inférence du quasi-maximum de vraisemblance de
modèles de séries chronologiques
à valeurs entières.**

Soutenue, publiquement, le 16 /04/2018 à 9h : 30, devant le Jury composé de :

M. M. BENTARZI	Professeur	à U.S.T.H.B :	Président.
M. A. AKNOUCHE	Professeur	à U.S.T.H.B :	Dir de thèse.
M. Y. Berkoun	Professeur	à Univ. de Tizi Ouzou	Examineur.
M. A. BIBI	Professeur	à Univ. de Mentouri	Examineur.
M. K. KHALED	Professeur	Constantine	Examineur.
M. M. MILOUD	Professeur	à Univ de Boumerdes	Examineur.
M. K. TOUFIK	Maître de conférence /A.	à U.S.T.H.B :	Examineur.
M. M. Maamra	Maître de conférence /B.	à U.S.T.H.B :	Invité.
		à U.S.T.H.B :	

Table des matières

Résumé	iv
Abstract	vi
Remerciements	viii
Introduction générale	xi
1 Modèle <i>INGAR</i>	1
1.1 Introduction	1
1.2 Traits bien connus des séries chronologiques à valeurs entières observées en pratique	2
1.3 Modèles basés sur la régression discrète	6
1.3.1 Modèle <i>INGARCH</i> poissonien	6
1.3.2 Modèle <i>INGARCH</i> binomial négatif _{<i>K</i>}	10
1.3.3 Modèle <i>INGARCH</i> log-linéaire	13
1.4 Modèles basés sur l'opérateur d'amincissement	16
1.4.1 Modèle <i>INAR</i> (1)	17
1.4.2 Modèle <i>INAR</i> (<i>p</i>)	18
1.4.3 D'autres généralisations	19
1.5 Approche <i>GLM</i> (generalized linear models) des modèles de séries chronologiques à valeurs entières	21

1.6	Généralisation : le modèle <i>INGAR</i> (INteger Generalized AutoRegression) . . .	22
2	Estimateur du <i>PQML</i> du modèle <i>INGAR</i>	25
2.1	Introduction	25
2.2	Estimateur du <i>QML</i> poissonien (<i>PQML</i>) du modèle <i>INGAR</i>	25
2.3	Propriétés asymptotiques de l'estimateur du <i>PQML</i>	26
2.4	Applications à des modèles de séries chronologiques à valeurs entières spécifiques	28
2.4.1	Application au modèle <i>INGARCH</i> poissonien	29
2.4.2	Application au modèle <i>INGARCH</i> binomial négatif _{κ}	30
2.4.3	Application au modèle <i>INAR</i> (1)	32
3	Estimateur du <i>GQML</i> du modèle <i>INGAR</i>	35
3.1	Introduction	35
3.2	Estimateur du <i>QML</i> géométrique (<i>GQML</i>) du modèle <i>INGAR</i>	35
3.3	Propriétés asymptotiques de l'estimateur du <i>GQML</i>	36
3.4	Applications à des modèles à valeurs entières spécifiques	38
3.4.1	Application au modèle <i>INGARCH</i> poissonien	38
3.4.2	Application au modèle <i>INGARCH</i> binomial négatif ₂	39
3.4.3	Application au modèle <i>INAR</i> (1)	40
3.5	Preuves des principaux résultats	43
4	Estimateur du <i>NB-QML</i> du modèle <i>INGAR</i>	46
4.1	Introduction	46
4.2	Estimateur du <i>QML</i> binomial négatif en profil (<i>pNB-QML</i>)	48
4.3	Propriétés asymptotiques du <i>QMLE</i> binomial négatif en profil	49
4.4	Application à des modèles spécifiques	52
4.4.1	Application au modèle <i>INGARCH</i> poissonien	52

4.4.2	Application au modèle <i>INGARCH</i> binomial négatif ₂	53
4.4.3	Application au modèle <i>INAR</i> (1)	54
4.5	Estimation en quatre étapes : estimateur du <i>QML</i> binomial négatif en deux étapes (<i>2SNB-QML</i>)	57
4.6	Propriétés asymptotiques du <i>2SNB-QMLE</i>	58
4.7	Application au modèle <i>INGARCH</i> binomial négatif ₂	60
4.8	Preuves des principaux résultats	61
5	Comparaison des estimateurs du QML	68
5.1	Introduction	68
5.2	Comparaison sous des modèles <i>INGARCH</i> (1, 1)	69
5.2.1	Comparaison vis-à-vis de l'efficacité asymptotique relative	69
5.2.2	Comparaison en échantillons finis : étude de simulation	71
5.3	Comparaison sous le modèle <i>INAR</i> (1)	80
5.3.1	Comparaison vis-à-vis de l'efficacité asymptotique relative	80
5.3.2	Comparaison en échantillons finis : étude de simulation	81
6	Applications réelles	83
6.1	Introduction	83
6.2	Application à la série "Polio"	84
6.3	Application à la série "transaction"	90
7	Conclusion générale	94
8	Annexe A	96
8.1	Glossaire	96
8.2	Liste des Tableaux	99
8.3	Liste des Figures	100

Résumé

La présente thèse porte sur l'inférence du quasi-maximum de vraisemblance (QML) d'une classe générale de modèles de séries chronologiques à valeurs entières. Cette classe de modèles que nous appellerons « $INGAR$ » (INteger Generalized AutoregRession) est spécifiée à travers la moyenne conditionnelle qui est une fonction générale, mesurable et paramétrée du passé infini du processus observable. Une telle classe, dans laquelle seule la moyenne conditionnelle est à spécifier, est très générale tant par la richesse des distributions conditionnelles offertes que par la forme générale de la moyenne conditionnelle qui peut être linéaire ou non-linéaire. Elle englobe en fait une large gamme de modèles de séries chronologiques à valeurs entières, notamment le modèle $ARMA$ généralisé ($GARMA$), le modèle $GARCH$ à valeurs entières ($INGARCH$) et le modèle $INAR$ (Integer AutoregRession). Pour cette classe, nous proposons deux variantes de l'estimateur du QML binomial négatif ($NB-QMLE$) calculées sur la base du profil de la vraisemblance binomiale négative. La première, appelée « $NB-QMLE$ en profil » ($pNB-QMLE$), consiste à maximiser la fonction de vraisemblance binomiale négative par rapport au paramètre de moyenne conditionnelle tout en fixant arbitrairement le paramètre de dispersion correspondant. Lorsque ce dernier est égal à l'unité, l'estimateur résultant se réduit à l'estimateur du QML géométrique. Le deuxième estimateur, consiste en quatre étapes : un estimateur du $pNB-QML$ en deux étapes pour estimer le paramètre de moyenne conditionnelle et un estimateur des moindres carrés pondérés en deux étapes pour estimer le paramètre de dispersion. Pour cela, le modèle $INGAR$ doit satisfaire une fonction de lien GLM (Generalized Linear Models) binomiale négative impliquant le paramètre de dispersion à estimer. Sous des hypothèses assez faibles, nous montrons que les deux estimateurs proposés sont consistants et asymptotiquement gaussiens sans avoir à spécifier entièrement la distribution conditionnelle du modèle. De plus, sous la fonction de lien GLM binomiale négative, l'estimateur en quatre étapes est asymptotiquement efficace

dans la classe de tous les estimateurs du QML appartenant à la famille exponentielle linéaire, y compris l'estimateur du QML poissonien. Des applications à des séries simulées et réelles montrent la performance des estimateurs proposés que nous comparons avec l'estimateur du QML poissonien.

Mots-clés et phrases : Modèles de séries chronologiques à valeurs entières, modèle $GARCH$ à valeurs entières, modèle $INAR$, Modèles Linéaires Généralisés, Quasi-vraisemblance, $QMLE$ géométrique, $QMLE$ binomial négatif, $QMLE$ poissonien, consistance et normalité asymptotique.

Abstract

This thesis concerns the quasi-maximum likelihood (QML) inference for a general class of integer valued time-series models. This class of models that we will call $INGAR$ (Integer Generalized Autoregression) is specified through the conditional mean which is a measurable, nonlinear and parametric function of the infinite past of the observable process. Such a class, in which only the conditional mean is to be specified, is very general both by the richness of the conditional distributions offered and also by the general form of the conditional mean which may be linear or nonlinear. It encompasses a wide range of integer time series models including the generalized $ARMA$ model ($GARMA$), the Integer-valued $GARCH$ model ($INGARCH$), and the $INAR$ (Integer Autoregression) model. For this class, we propose two variants of the negative binomial $QMLE$ ($NB-QMLE$) calculated on the basis of the profile of the negative binomial likelihood. The first one, called "profile $NB-QMLE$ " ($pNB-QMLE$), consists in maximizing the negative binomial likelihood function with respect to the conditional mean parameter while arbitrarily fixing the corresponding dispersion parameter. When the latter is equal to unity, the resulting estimate reduces to the geometric $QMLE$. The second estimate, however, consists of four stages : a two-stage $pNB-QML$ estimator to estimate the conditional mean parameter and a two-stage weighted least-squares estimator to estimate the dispersion parameter. To do this, the $INGAR$ model should satisfy a negative binomial GLM (Generalized Linear Models) link function involving the dispersion parameter to be estimated. Under fairly weak assumptions, we show that the two proposed estimates are consistent and asymptotically Gaussian without having to fully specify the conditional distribution of the model. Moreover, under the negative binomial GLM link function, the four-stage estimate is asymptotically efficient in the class of all $QMLEs$ belonging to the linear exponential family, including the Poisson $QMLE$. Applications to simulated and real series show the performance of the proposed estimators

which we compare among others with the Poisson *QMLE*.

Keywords and phrases : Integer-valued time series models, Integer *GARCH*, Integer *AR*, Generalized Linear Models, Quasi-likelihood, geometric *QMLE*, negative binomial *QMLE*, poisson *QMLE*, consistency and asymptotic normality.

Remerciements

D'abord, j'adresse mes remerciements à mon directeur de thèse, le Professeur Abdelhakim Aknouche, qui m'a dirigée tout au long de cette période d'étude. **A**insi, je lui suis très reconnaissante pour son assertion, son immédiateté et son luxuriance dans ses réponses. **O**utre ses qualités humaines, je n'oublierai jamais son rôle majeur en tant que professeur. D'éveiller la joie de travail et de connaissances me créant une extraordinaire sensation de sécurité et de paix absolue. **C**ette expérience était très importante pour ma formation ainsi que pour les tâches auxquelles il m'a associée, lesquelles m'ont vraiment permises de consolider mes connaissances et d'en développer de nouvelles. **H**umblement, je voudrais lui dire merci pour son soutien pendant mes périodes de doutes et ses multiples encouragements répétés. **E**nfin, être une de ses étudiantes m'a permise non seulement de tracer mon chemin mais de mesurer l'ampleur des problèmes scientifiques auxquels j'allais faire face.

Je tiens à remercier Monsieur le Professeur Mohamed Bentarzi pour l'honneur qu'il m'a fait en acceptant d'être président de mon jury de thèse. Je tiens à l'assurer de ma profonde reconnaissance pour l'intérêt qu'il porte à ce travail.

J'exprime tous mes remerciements à l'ensemble des membres du jury. Je suis honorée par la présence des Professeurs Abdelouhab Bibi, Khaled Khaldi, Miloud Mihoubi, Youcef Berkoun et Toufik Kernane. Je remercie de plus Monsieur le Docteur Mohamed Said Maamra pour l'intérêt qu'il a manifesté en participant en qualité de membre invité à ce jury. C'est un honneur pour moi de pouvoir compter sur leurs lectures et leurs critiques.

Je suis infiniment gré à Monsieur Rabah Messaci et Yasmina Messaci pour leur aide précieuse et leurs soutiens scientifiques notamment dans l'outil de programmation, le logiciel R.

J'adresse toute ma gratitude à tous mes ami(e)s et à toutes les personnes qui m'ont aidé dans la réalisation de ce travail.

Enfin, les mots les plus simples étant les plus forts, j'adresse toute mon affection à ma famille, et en particulier à ma maman pour son soutien qui m'a été bien utile durant ma thèse. A titre plus personnel, je remercie chaleureusement mon mari, Abderrahmane, pour la grande patience, l'encouragement et la confiance qu'il m'a témoignée. Je tiens à le remercier surtout pour son soutien moral ininterrompu et ses nombreux conseils tout le long de ma thèse.

A la mémoire de mon Père et à mon fils Adem.

Introduction générale

L'analyse classique des séries chronologiques a été quasiment basée sur des modèles que sont des équations aux différences stochastiques dont les entrées et les formes sont telles que les termes des processus solutions (ou sorties) soient absolument continues (notamment ayant une loi normale, loi de Student, mélange de lois absolument continues...). Pourtant, il est des séries chronologiques à valeurs entières, telles les données de comptage (count data) et les données binaires, qui ne peuvent être représentées adéquatement par ces modèles classiques, sous peine d'une imprécision tant dans la modélisation que dans l'exploitation (prédiction, filtrage...) de celle-ci. C'est pourquoi, un intérêt considérable a été accordé au cours des trois dernières décennies à des modèles de séries chronologiques à valeurs entières dont la distribution du processus sortie conditionnellement à son passé est plutôt discrète (ex. loi de Poisson, loi binomiale négative, loi binomiale...).

De nombreux modèles de séries temporelles à valeurs entières ont été récemment introduits à tel point qu'il est difficile de les classer. Cependant, deux classes majeures de modèles ont joué un rôle central. La première est la classe de modèles basés sur la régression discrète tels que les modèles *ARMA généralisés* (*GARMA*), l'autorégression de Poisson et en particulier le fameux modèle *autorégressif conditionnellement hétéroscédastique généralisé à valeurs entières* (*INGARCH*) (Benjamin et al, 2003; Heinen, 2003; Ferland et al, 2006; Fokianos et al, 2009; Zhu, 2011-2012a-2012b-2012c; Doukhan et al, 2012; Christou et Fokianos, 2014; Davis et Liu, 2016; Chen et al, 2016). Quant à la seconde classe, elle concerne les modèles basés sur les équations aux différences stochastiques faisant intervenir l'opérateur d'amincissement dont l'exemple typique est le modèle *AR entier* (*INAR*) (Integer Autoregression) (par exemple. McKenzie, 1985-2003; Al-Osh et Alzaid, 1987; Silva, 2015; Bourguignon, 2016).

Ahmad et Francq (2016) ont récemment introduit un modèle plus général englobant les deux classes susmentionnées. Ce modèle que nous appelons *INteger Generalized Auto-Regression* (désormais *INGAR*) est défini en spécifiant sa moyenne conditionnelle comme étant une fonction mesurable et paramétrée du passé infini du processus observable. Des sous-classes importantes de ce modèle sont *l'auto-régression de Poisson* générale (Doukhan et al, 2012; Doukhan et Kengne, 2015; Kengne, 2015), le modèle *INGARCH* et le modèle *INAR*. Pour le modèle *INGAR*, Ahmad et Francq (2016) ont établi la consistance et la normalité asymptotique de *l'estimateur du quasi-maximum de vraisemblance poissonien* (*P-QMLE*) qui est calculé comme si la distribution conditionnelle des observations était poissonienne. Le *P-QMLE* a en effet de nombreuses qualités statistiques : *i*) premièrement, il est robuste devant une mauvaise spécification de la vraie loi conditionnelle du modèle pourvu que la moyenne conditionnelle soit bien spécifiée. Cela est dû au fait que la fonction de vraisemblance poissonienne appartient à la famille exponentielle linéaire (White, 1982; Gourieroux et al, 1984a-1984b). *ii*) deuxièmement, il est asymptotiquement efficace lorsque la vraie loi conditionnelle des données est poissonienne. *iii*) troisièmement, lorsque la variance et la moyenne conditionnelles du modèle sont proportionnelles, le *P-QMLE* est asymptotiquement efficace dans la classe de tous les *QMLEs* dont la fonction de vraisemblance appartient à la famille exponentielle linéaire (voir Gourieroux et al, 1984a). Cette dernière proportionnalité entre moyenne et variance conditionnelles est habituellement appelée hypothèse (ou fonction de lien) du *Modèle Linéaire Généralisé* (désormais *GLM*) poissonien. Cependant, malgré ces avantages, la distribution de Poisson, qui est connue pour être équidispersée, s'adapte difficilement aux séries surdispersées fréquemment observées dans la pratique. Par conséquent, fort probablement le *P-QMLE* n'atteindrait pas sa pleine efficacité asymptotique en présence de données surdispersées. Ainsi, le recours à un *estimateur du quasi-maximum de vraisemblance (QMLE)*, calculé à l'aide d'une fonction de vraisemblance surdispersée tout en appartenant à la famille exponentielle linéaire, pourrait être un complément intéressant au *P-QMLE*.¹

¹Largement inspirée de Aknouche (2013).

Pour le modèle *INGAR* considéré par Ahmad et Francq (2016), nous proposons dans cette thèse deux estimateurs du *quasi-maximum de vraisemblance* binomial négatif (*NB-QMLE*). Ces estimateurs sont calculés sur la base d'un profil de la fonction de vraisemblance binomiale négative, appartenant à la famille exponentielle linéaire.

Le premier, que nous appellerons «*NB-QMLE* en profil» (*pNB-QMLE*), consiste à maximiser la fonction de vraisemblance binomiale négative par rapport au paramètre de moyenne conditionnelle tout en gardant arbitrairement fixé le paramètre de dispersion correspondant. En particulier, lorsque ce dernier est égal à 1, l'estimateur résultant se réduit au *QMLE géométrique* (Aknouche et Bendjedou, 2017).

Le deuxième, cependant, se compose de quatre étapes : un *NB-QMLE* en deux étapes pour estimer le paramètre de moyenne conditionnelle du modèle et un estimateur des moindres carrés pondérés en deux étapes pour le paramètre de dispersion. Pour cela, le modèle *INGAR* devrait satisfaire *une fonction de lien GLM* binomiale négative impliquant le paramètre de dispersion inconnu à estimer. Dans le contexte de la régression statique à valeurs entières, un estimateur similaire en trois étapes a été appelé «*estimateur du pseudo-maximum de vraisemblance quasi-généralisée*» par Gourieroux et al (1984b) et «*estimateur du quasi-maximum de vraisemblance binomial négatif en deux étapes*» (*2SNB-QMLE*) par Wooldridge (1997, 2002). Adoptant cette dernière notation, l'estimateur en quatre étapes que nous proposons sera désigné par l'acronyme *2SNB-QMLE*.

Nous montrons sous des hypothèses assez faibles et générales que les deux estimateurs proposés sont fortement consistants et asymptotiquement gaussiens. De plus, sous la fonction de lien *GLM* binomiale négative, le *2SNB-QMLE* est asymptotiquement efficace dans la classe de tous les *QMLEs* appartenant à la famille exponentielle linéaire, y compris le *P-QMLE*. Nous comparons ensuite nos deux estimateurs proposés avec l'estimateur du *QML* poissonien asymptotiquement et en échantillons finis. D'abord nous comparons l'efficacité relative des deux estimateurs proposés avec celle du *P-QMLE* sur plusieurs classes importantes du modèle *INGAR*, notamment le modèle *INGARCH* poissonien, le modèle *INGARCH* binomial négatif de type 1, le modèle *INGARCH* binomial négatif de type 2 et le modèle *INAR(1)*. Pour des modèles *INGARCH* avec distributions conditionnelles double Poisson et Poisson généralisée voir Aknouche et al (2018). Ensuite, nous comparons

les performances des estimateurs P - $QMLE$, pNB - $QMLE$ et $2SNB$ - $QMLE$ sur des séries simulées des modèles $INGARCH$ et $INAR$ mentionnés. Enfin, dans le cadre du modèle $INGARCH$ binomial négatif de type 2, nous proposons d'appliquer le $2SNB$ - $QMLE$ sur deux séries chronologiques à valeurs entières bien connues dans la pratique. La première connue sous le nom de "données Polio" représente le nombre mensuel de cas de poliomyélite aux États-Unis sur la période allant de 1970 à 1983 (données Polio, Zeger, 1988; Zeger et Qaqish, 1988; Davis et *al*, 1999; Benjamin et *al*, 2003; Heinen, 2003; Davis et Wu, 2009; Zhu, 2011). La seconde appelée "données transaction" est le nombre de transactions par minute pour le stockes Ericsson B au courant du 05 Juillet, 2002 (e.g. Fokianos et *al*, 2009; Davis et Liu, 2016; Christou et Fokianos, 2014).

Nous présentons maintenant avec plus de détails le contenu de la thèse ainsi que la méthodologie suivie lors de notre recherche.

Chapitre 01 : Modèle $INGAR$ (INteger Generalized AutoRegression) : une classe générale de modèles de séries chronologiques à valeurs entières

Dans ce chapitre, nous passons en revue quelques aspects marquants des séries chronologiques à valeurs entières en donnant quelques exemples de simulations. Ensuite, nous examinons deux classes générales de ces modèles via une étude probabiliste : les modèles basés sur la régression discrète dont l'exemple le plus représentatif est le modèle $INGARCH$, et les modèles basés sur les équations aux différences stochastiques faisant intervenir l'opérateur d'amincissement (thinning operator) dont l'exemple typique est le modèle $INAR$. Par la suite nous mettons l'accent sur une approche considérée comme étant un moyen permettant d'unifier d'autres modèles statistiques qui est l'approche des modèles linéaires généralisés (GLM, Generalized Linear Model). En particulier, la fonction de lien GLM liant la moyenne conditionnelle et la variance conditionnelle du modèle est exhibée. Nous terminerons ce chapitre par la présentation d'un modèle, noté $INGAR$, généralisant les deux classes précédentes en donnant sa structure, son utilité majeure et en identifiant les deux cas considérés par ce type de modèles.

Chapitre 02 : Estimateur du quasi-maximum de vraisemblance poissonien du modèle *INGAR*

Ce chapitre cerne et synthétise les principaux résultats de Ahmad et Francq (2016) qui ont proposé un *estimateur du quasi-maximum de vraisemblance poissonien (P-QMLE)* pour le modèle *INGAR*. Dans la première section, nous exploitons la structure de l'estimateur en question tandis que la deuxième section met l'accent sur la consistance et la normalité asymptotique de l'estimateur du *QML* poissonien. Enfin, dans la dernière section, nous explicitons les résultats précédents à travers des applications sur des modèles à valeurs entières spécifiques tels que le modèle *INGARCH* (1, 1) poissonien, le modèle *INGARCH* (1, 1) binomial négatif de type 2 ainsi que le modèle *INAR*(1).

Chapitre 03 : Estimateur du quasi-maximum de vraisemblance géométrique (*G-QML*) du modèle *INGAR*

Dans ce chapitre nous proposons un estimateur du quasi-maximum de vraisemblance géométrique (*GQMLE*) calculé comme si la distribution conditionnelle du processus observable était géométrique. Dans la première section, nous définissons l'estimateur sous-jacent puis nous étudions ses propriétés asymptotiques sous un jeu d'hypothèses assez faibles ; à savoir la consistance et la normalité asymptotique qui font l'objet d'étude de la section suivante. Dans la troisième section, nous donnons des exemples clarifiant les résultats liés à l'inférence statistique de l'estimateur en question.

Chapitre 04 : Estimation du quasi-maximum de vraisemblance binomial négatif (NB-QML) du modèle *INGAR*

Ce chapitre est considéré comme étant le cœur de cette thèse vu que son objet contient deux nouveaux estimateurs proposés dans le cadre de ce travail. En guise d'introduction, nous présentons le critère du maximum de vraisemblance binomial négatif tout en soulignant ses inconvénients majeurs. Nous présentons ensuite la fonction du "*Quasi-Vraisemblance binomiale négative en Profil*" à partir de laquelle sont définis les deux estimateurs proposés dans ce chapitre. Dans la Section 1, considérant le **1^{er} cas** du modèle *INGAR* où seule la moyenne conditionnelle est spécifiée, nous proposons l'*Estimateur du Quasi Maximum de Vraisemblance binomial négatif en profil*, *pNB-QMLE*. La Section 2 établit les propriétés asymptotiques de cet estimateur tandis que la troisième Section complète la précédente à travers des exemples. La Section 4 considère le **2^{ème} cas** du modèle *INGAR* où il existe deux paramètres à estimer ; le paramètre de moyenne conditionnelle du modèle (déjà estimé dans la section 1) et le paramètre de dispersion. En particulier, nous présentons le principe de la méthode d'estimation du paramètre de dispersion "*Moindres Carrés Pondérés*" ensuite nous résumons toutes les étapes à travers un algorithme que nous appelons "*NB-QMLE en deux étapes*". L'estimateur résultant sera appelé "*Estimateur du Quasi Maximum de Vraisemblance binomial Négatif en deux étapes*", *2SNB-QMLE*. En effet, deux étapes sont nécessaires pour estimer le paramètre de moyenne conditionnelle du modèle et deux autres sont pour estimer le paramètre de dispersion. La cinquième section établit la consistance et la normalité asymptotique du *2SNB-QMLE* alors que la sixième section applique les résultats asymptotiques dans le cadre du modèle *INGARCH* binomial négatif de type 2. Enfin, la dernière section offre toutes les démonstrations des principaux résultats utilisés dans ce chapitre.

Chapitre 05 : Comparaison des méthodes du QML dans le cadre du modèle $INGAR$

Ce chapitre repose sur une vision comparative des estimateurs proposés avec le P-QMLE. Cette comparaison prend deux volets. D'abord vis-à-vis de l'efficacité asymptotique relative des estimateurs, ensuite en échantillons finis à travers une étude de simulation. Ainsi ce chapitre est partagé principalement en deux sections. En outre, la section 1 compare, sous certaines hypothèses GLM spécifiques, les variances asymptotiques des estimateurs du QML binomiaux négatifs proposés avec celle de l'estimateur du QML poissonien. La comparaison est faite par rapport à trois instances du modèle $INGAR$ à savoir le modèle $INGARCH(1, 1)$ poissonien, le modèle $INGARCH(1, 1)$ géométrique et le modèle $INGARCH(1, 1)$ binomial négatif de types 1 et 2. La seconde section concerne une autre classe particulière du modèle $INGAR$ nommée « les modèles $INAR(1)$ » dont l'enchaînement des étapes est le même que lors de la précédente section.

Chapitre 06 : Applications réelles

Ce chapitre illustre les résultats présentés dans les chapitres précédents à travers des applications réelles de notre méthode d'estimation $2SNB-QMLE$. Dans le cadre du modèle $INGARCH(1, 1)$ binomial négatif de type 2, nous considérons deux célèbres séries temporelles à valeurs entières. La première est le nombre mensuel de cas de poliomyélite aux États-Unis sur la période allant de 1970 à 1983 avec un total de $n = 168$ observations (données Polio, Zeger, 1988). La deuxième consiste en le nombre de transactions par minute pour le stock Ericsson B au courant du 05 Juillet, 2002 (données transaction, Fokianos et al, 2009).

Chapitre 1

Modèle *INGAR* (INteger Generalized AutoRegression) : une classe générale de modèles de séries chronologiques à valeurs entières

1.1 Introduction

Dans ce chapitre, nous présentons à travers des simulations certains faits stylisés caractérisant les séries chronologiques à valeurs entières. Ensuite, nous nous intéressons principalement à étudier quelques propriétés probabilistes de deux cas particuliers de modèles de séries chronologiques à valeurs entières ; à savoir le modèle *INGARCH* et le modèle *INAR*. Par la suite nous mettrons l'accent sur une approche considérée comme étant un moyen permettant d'unifier d'autres modèles statistiques tels que la régression linéaire, la régression logistique et la régression de Poisson. Cette approche est connue sous le nom "Modèles Linéaires Généralisés", GLM. Nous terminerons ce chapitre par la présentation d'un modèle général englobant les deux cas précédents, noté *INGAR*, en donnant sa structure et en précisant son utilité majeure. Tout au long de ce chapitre, nous désignons par \mathcal{F}_t la σ -algèbre générée par $\{X_t, X_{t-1}, \dots\}$.

1.2 Traits bien connus des séries chronologiques à valeurs entières observées en pratique

L'analyse statistique de certaines données discrètes présente une variabilité plus grande que la moyenne de l'échantillon. Ce phénomène dit de *surdispersion* a été largement et diversement étudié dans la littérature, en particulier en relation avec la loi de Poisson. Si un modèle inadéquat est adopté en présence de la surdispersion, il peut y avoir perte d'efficacité pour les différentes statistiques en jeu. Cox (1983) a étudié en détail cet effet pour le modèle poissonien. Outre cette particularité, les séries chronologiques à valeurs entières observées en pratique, notamment les séries de comptage, sont caractérisées par des traits spécifiques tels que la présence importante de petites valeurs, la sur-fréquence de zéros, l'asymétrie, une structure d'autocorrélation positive et un comportement localement constant. Pour inspecter et visualiser ces caractéristiques, nous utiliserons les fichiers de données Polio "nombre mensuel de cas de poliomyélite aux Etats-Unis sur la période allant de 1970 à 1983" et "nombre de transactions par minute pour le stockes Ericsson B au cours de 05 Juillet, 2002" dont les valeurs sont indiquées dans les tableaux suivants :

0	1	0	0	1	3	9	2	3	5	3	5	2	2	0	1	0	1	3	3	2	1	1	5
0	3	1	0	1	4	0	0	1	6	14	1	1	0	0	1	1	1	1	0	1	0	1	0
1	0	1	0	1	0	1	0	1	0	0	2	0	1	0	1	0	0	1	2	0	0	1	2
0	3	1	1	0	2	0	4	0	2	1	1	1	1	0	1	1	0	2	1	3	1	2	4
0	0	0	1	0	1	0	2	2	4	2	3	3	0	0	2	7	8	2	4	1	1	2	4
0	1	1	1	3	0	0	0	0	1	0	1	1	0	0	0	0	0	1	2	0	2	0	0
0	1	0	1	0	1	0	2	0	0	1	2	0	1	0	0	0	1	2	1	0	1	3	6

Tableau 1.2.1. Données de la série Polio, $n = 168$ observations.

18	9	13	24	17	13	18	9	19	8	9	10	11	17	11	5	8	10	11
8	9	22	13	11	8	17	8	16	10	15	17	16	11	6	5	13	13	11
11	12	14	16	16	11	10	11	8	12	16	14	10	7	10	4	12	8	7
13	9	5	22	11	12	13	8	9	3	7	13	21	21	15	12	10	21	9
12	5	14	11	16	33	26	10	19	14	20	17	12	16	17	17	7	17	11
14	11	11	13	18	7	4	12	12	6	21	19	7	12	10	6	12	7	5
11	10	13	17	9	10	13	16	11	12	16	12	21	10	8	6	7	7	12
17	10	10	3	14	14	13	13	7	4	8	9	6	17	5	2	6	1	7
6	6	8	7	8	13	9	11	5	7	11	0	10	9	6	6	3	3	6
7	9	4	8	18	7	11	15	9	9	9	17	12	6	13	8	22	14	8
13	13	8	14	10	3	8	12	9	4	12	9	7	9	15	7	7	5	16
9	8	8	8	7	13	11	5	11	16	7	7	8	2	10	10	9	4	9
12	4	12	11	7	10	11	9	7	9	10	4	5	7	10	13	12	7	8
8	12	12	9	12	10	9	10	7	12	4	8	14	7	10	15	6	6	2

3	7	8	6	6	5	6	6	6	14	6	7	31	10	14	8	17	7	9
11	13	6	7	11	7	7	5	6	11	24	12	10	8	11	16	12	13	5
6	11	4	14	11	7	14	8	12	10	9	3	11	10	7	16	10	9	2
11	10	5	10	8	8	1	6	13	8	3	1	9	9	6	4	7	1	9
4	14	6	12	7	1	6	5	3	3	8	6	3	12	12	21	9	5	20
20	10	5	8	13	6	5	7	3	5	8	9	11	16	24	28	6	20	8
9	5	16	11	5	8	12	19	7	10	5	7	6	6	5	11	14	10	7
8	6	10	5	3	6	12	7	3	8	3	14	4	6	5	2	3	13	5
10	5	9	17	4	9	10	8	6	6	14	7	4	8	14	9	9	12	5
1	3	5	7	11	21	7	5	4	8	18	4	8	3	10	13	8	4	11
8	13	9	9															

Tableau 1.2.2. Données de la série nombre de transactions, $n = 460$ observations.

Des tableaux précédents, on peut vérifier facilement que la condition de surdispersion est trivialement vérifiée vu que la variance empirique excède la moyenne empirique pour les deux cas. En effet, la série Polio est clairement surdispersée vu que sa moyenne échantillonnale est égale à 1,3333 et sa variance échantillonnale vaut 3,5050. A l'instar des données Polio, la série transaction est également surdispersée dont la moyenne et variance échantillonnales sont respectivement égales à 9.8239 et 23.7532. Maintenant, en ce qui concerne les autres spécificités, elles sont exhibées dans les figures ci-dessus.

- o Présence importante de petites valeurs :

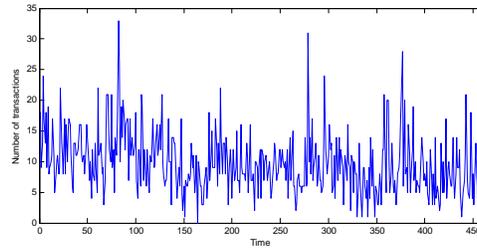


Figure 1.2.1 : Graphe de la série transaction.

- o Excès de zéros et Comportement localement constant¹ :

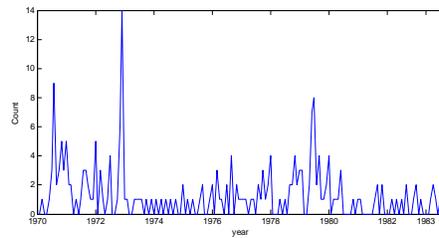


Figure 1.2.2 : Graphe de la série Polio.

¹voir également Zeger, 1988, Benjamin et al, 2003, Zhu, 2011

○ Asymétrie :

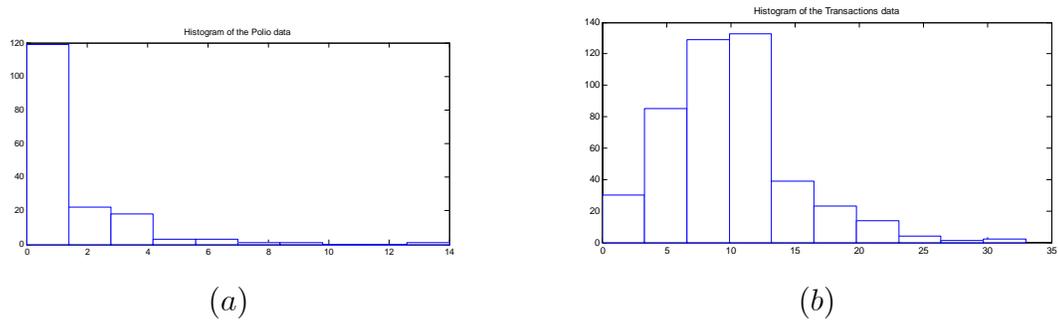


Figure 1.2.3 : Histogrammes des deux séries.
 (a) Série Polio. (b) Série transaction.

○ Structure de corrélation positive :

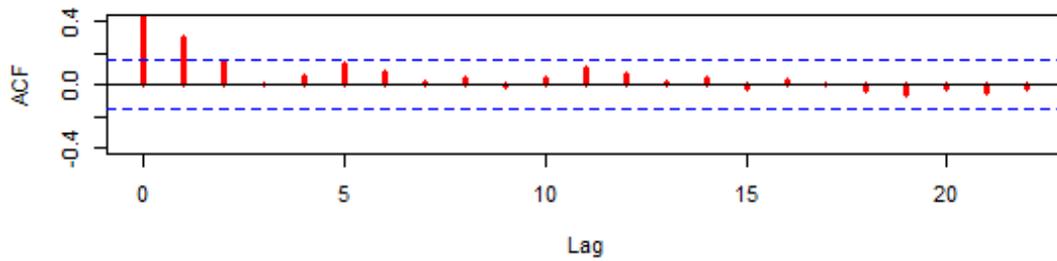


Figure 1.2.4 : Corrélogramme de la série Polio.

Il est bien connu qu'en raison de ces caractéristiques, les séries chronologiques à valeurs entières ne peuvent en général être représentées adéquatement par des modèles de séries chronologiques à valeurs continues, notamment les modèles de type ARMA, GARCH,...

1.3 Modèles basés sur la régression discrète

Cette section a comme but d'étudier un certain nombre de modèles qui ont été proposés pour les séries chronologiques à valeurs discrètes, tout en soulignant quelques aspects probabilistes tels que les distributions marginales, les conditions de stabilité et la structure de corrélation. Un point de départ naturel est de commencer par l'outil de base populaire "le modèle de régression poissonien", qui est en effet, le modèle de base de représentation des données de séries de comptage. Ensuite nous présentons une autre hypothèse de répartition, la distribution binomiale négative et nous terminons par un autre type de ces modèles, connu par "les modèles log-linéaires".

1.3.1 Modèle *INGARCH* poissonien

Un modèle couramment utilisé pour les séries temporelles de comptage avec une surdispersion est le modèle autorégressif conditionnellement hétéroscédastique généralisé à valeurs entières (*INGARCH*) dont la distribution conditionnelle est supposée être poissonienne. Un processus *GARCH* poissonien $\{X_t, t \in \mathbb{Z}\}$ à valeurs entières (*INGARCH* (p, q)), comme proposé par Rydberg et Shephard (2000) et Grunwald et *al.* (2000), est défini par une distribution conditionnelle poissonienne

$$X_t \mid \mathcal{F}_{t-1} \sim \mathcal{P}(\lambda_t), \quad t \in \mathbb{Z}, \quad (1.3.1a)$$

avec une moyenne conditionnelle $\lambda_t = \lambda_t(\theta_0)$ spécifiée comme suit

$$\lambda_t(\theta_0) = \omega_0 + \sum_{i=1}^q \alpha_{0i} X_{t-i} + \sum_{j=1}^p \beta_{0j} \lambda_{t-j}(\theta_0), \quad (1.3.1b)$$

où le paramètre inconnu $\theta_0 = (\omega_0, \alpha_{01}, \dots, \alpha_{0q}, \beta_{01}, \dots, \beta_{0p})' \in \Theta \subset \mathbb{R}^m$ ($m \in \mathbb{N}^* = \mathbb{N} - \{0\}$) est tel que $\omega_0 > 0$, $\alpha_{0i} \geq 0$, $\beta_{0j} \geq 0$. Il est supposé aussi que X_t est non dégénéré et que, si $p > 0$, $A_{\theta_0}(z) = \sum_{i=1}^q \alpha_{0i} z^i$ et $B_{\theta_0}(z) = 1 - \sum_{i=1}^p \beta_{0i} z^i$ n'ont aucune racine en commun (voir Ahmad et Francq, 2016).

Ferland et al (2006) ont montré que sous la condition de stabilité suivante

$$\sum_{i=1}^q \alpha_{0i} + \sum_{j=1}^p \beta_{0j} < 1, \quad (1.3.2)$$

le processus $\{X_t, t \in \mathbb{Z}\}$ donné par (1.3.1) est strictement stationnaire (voir aussi Franke, 2010). L'ergodicité du modèle $INGARCH(p, q)$, (1.3.1), a été établie d'abord par Grunwald et al. (2000) pour $p = 0$, Fokianos et al. (2009), Neumann (2011), Davis et Liu (2016), Douc et al. (2013) pour le cas $p = q = 1$, et par Doukhan et al. (2012) et Gonçalves et al. (2015) pour p, q quelconques.

Pour mieux voir le comportement de ce modèle, nous avons simulé un processus $INGARCH(5, 2)$ avec les paramètres : $\omega_0 = 1$, $\beta_{0j} = (0.3, 0.2, 0.1, 0, 0.2)'$ et $\alpha_{0i} = (0.05, 0.05)'$, et comme le vecteur β_{0j} est 5-dimensionnel, nous devons donner les 5 premières valeurs de λ , i.e, $\lambda_0, \dots, \lambda_4$. Nous avons sélectionné $\lambda_i = 1, \forall i = 0, \dots, 4$. La figure 1.3.1 présente l'histogramme de 10.000 valeurs simulées auprès le modèle (1.3.1) suivie par la fonction densité poissonnienne en rouge.

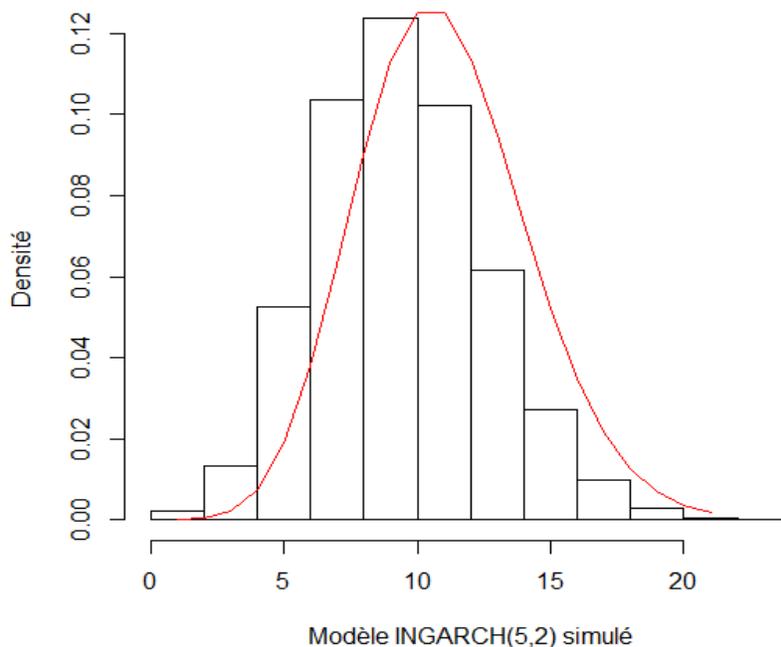


Figure 1.3.1 : Histogramme du processus INGARCH(5,2) simulé avec la fonction densité poissonnienne en rouge.

Cette dernière exhibe que le processus $INGARCH(5, 2)$ suit une loi poissonnienne qui tend vers la Normale. Dans notre cas, la valeur moyenne de λ est autour de 10. Ainsi, la distribution marginale en rouge est proche de celle normale.

En raison de simplicité, nous considérons dorénavant le cas particulier $p = q = 1$ et nous étudions les propriétés probabilistes de ce modèle. Pour ce faire, nous considérons le changement du bruit $\varepsilon_t = X_t - \lambda_t$ pour lequel l'équation (1.3.1b) devient

$$X_t - \varepsilon_t = \omega_0 + (\alpha_{01} + \beta_{01})X_{t-1} - \beta_{01}\varepsilon_{t-1}.$$

Notons que la suite de variables aléatoires $\{\varepsilon_t\} = \{X_t - \lambda_t\}$ est considérée comme étant un processus bruit blanc faible vu que sa moyenne et sa variance sont des constantes ainsi que ses variables aléatoires sont non corrélées entre elles :

$$\begin{aligned} E(\varepsilon_t) &= E(X_t - \lambda_t) = E(E(X_t - \lambda_t | \mathcal{F}_{t-1})) = 0, \\ var(\varepsilon_t) &= var(E(\varepsilon_t | \mathcal{F}_{t-1})) + E(var(\varepsilon_t | \mathcal{F}_{t-1})) \\ &= E(var(X_t | \mathcal{F}_{t-1})) = E(\lambda_t) = \mu, \\ Cov(\varepsilon_t, \varepsilon_{t+h}) &= E(\varepsilon_t E(\varepsilon_{t+h} | \mathcal{F}_{t+h-1})) = 0. \end{aligned}$$

Ainsi, les moments du processus $INGARCH(1, 1)$ sont tous finis si et seulement si $0 < \alpha_{01} + \beta_{01} < 1$. Dans ce cas :

1. La valeur moyenne du processus est donnée par : $\mu = \frac{\omega_0}{1 - (\alpha_{01} + \beta_{01})}$.

2. Sa variance est :

$$var(X_t) = var(\varepsilon_t) + var(\lambda_t) = \frac{(1 - (\alpha_{01} + \beta_{01})^2 + \alpha_{01}^2)}{1 - (\alpha_{01} + \beta_{01})^2} \mu.$$

3. De plus, sa fonction d'autocovariance est :

$$\gamma_X(h) = (\alpha_{01} + \beta_{01})^{h-1} \frac{\alpha_{01}\mu(1 - \beta_{01}(\alpha_{01} + \beta_{01}))}{1 - (\alpha_{01} + \beta_{01})^2}, \quad \forall h \geq 1$$

Encore une autre fois, si nous revenons à l'outil de simulation, nous construisons 10.000 observations pour 4 processus $INGARCH(1, 1)$ avec les paramètres présentés dans le tableau

suisant où la dernière colonne indique les valeurs moyennes correspondantes aux 4 processus :

Processus	valeurs de ω_0	valeurs de β_{01}	valeurs de α_{01}	valeurs de μ
$\{X_t\}$	1	0.1	0.2	1.4
$\{Y_t\}$	2	0.3	0.3	5
$\{Z_t\}$	3	0.5	0.4	30
$\{K_t\}$	2	0.45	0.5	40

Tableau 1.3.1 : Différentes valeurs des paramètres $(\omega_0, \beta_{01}, \alpha_{01})$ et les moyennes correspondantes aux 4 processus *INGARCH*(1, 1).

Du tableau précédent, on conclut que contrairement aux distributions marginales des processus $\{X_t\}$ et $\{Y_t\}$, celles des processus $\{Z_t\}$ et $\{K_t\}$ tendent vers la loi normale vu que leurs moyennes correspondantes sont supérieures à 10. Autrement dit, pour ces deux processus, la distribution normale est une bonne approximation de celle poissonnienne. Cet effet, est confirmé graphiquement par le tracé des histogrammes des 4 processus (voir figure 1.3.2).

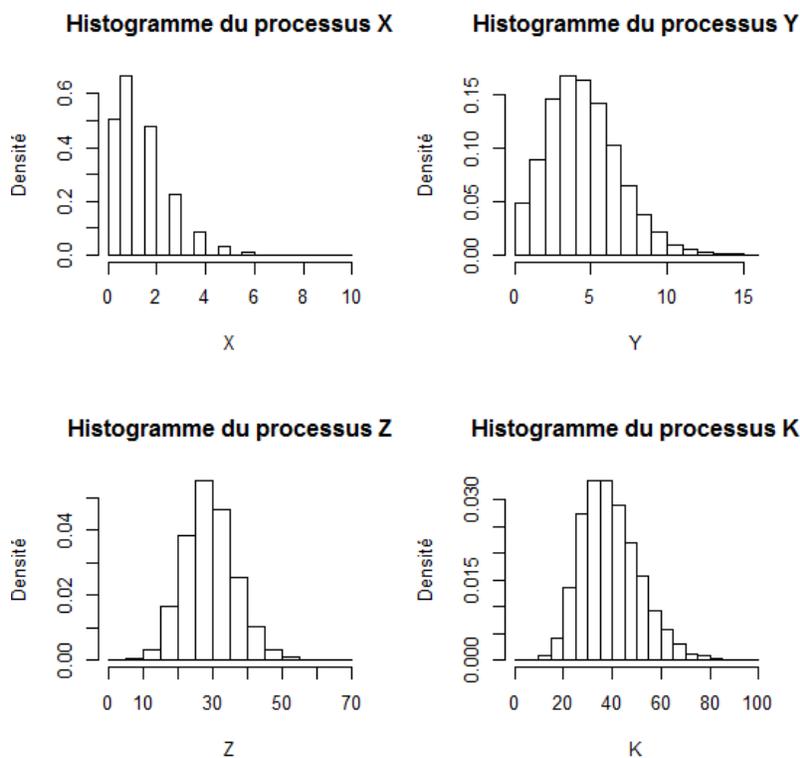


Figure 1.3.2 : Histogrammes des processus *INGARCH*(1,1) X_t, Y_t, Z_t et K_t respectivement.

1.3.2 Modèle *INGARCH* binomial négatif_K

La distribution de Poisson est le cas le plus naturel parmi les distributions discrètes pour des données de comptage. Cependant, la littérature offre plusieurs alternatives à celle-ci. Dans ce qui suit, nous discutons le cas d'une distribution binomiale négative_K² comme modèle alternatif pour l'analyse des séries de comptage.

Un processus $\{X_t, t \in \mathbb{Z}\}$ est dit *INGARCH* binomial négatif_K (*NB_K-INGARCH* (p, q)) si sa distribution conditionnelle est une binomiale négative,

$$X_t \mid \mathcal{F}_{t-1} \sim \mathcal{NB}(r_t, \pi_t), \quad t \in \mathbb{Z}, \quad (1.3.3a)$$

avec

$$r_t = r_0 \lambda_t^{2-K} \quad \text{et} \quad \pi_t = \frac{r_0 \lambda_t^{2-K}}{r_0 \lambda_t^{2-K} + \lambda_t}, \quad (1.3.3b)$$

où $K \in \mathbb{Z}$, $r_0 > 0$ et $\lambda_t = \lambda_t(\theta_0)$ satisfait la représentation *INGARCH* (p, q), (1.3.1b).

En considérant le modèle *NB₁-INGARCH* (p, q) correspondant à $K = 1$, i.e.

$$X_t \mid \mathcal{F}_{t-1} \sim \mathcal{NB}\left(r_0 \lambda_t, \frac{r_0 \lambda_t}{r_0 \lambda_t + \lambda_t}\right) \equiv \mathcal{NB}\left(r_0 \lambda_t, \frac{r_0}{r_0 + 1}\right), \quad (1.3.4)$$

ce modèle est strictement stationnaire et ergodique, avec un moment d'ordre deux fini, sous la même condition de stationnarité (1.3.2) pour le modèle *INGARCH* poissonien.

Considérons maintenant le modèle *NB₂-INGARCH* (p, q) qui correspond à (1.3.3) avec $K = 2$, i.e.

$$X_t \mid \mathcal{F}_{t-1} \sim \mathcal{NB}\left(r_0, \frac{r_0}{r_0 + \lambda_t}\right), \quad (1.3.5)$$

où $r_0 > 0$ et λ_t est donné par (1.3.1b). Le modèle (1.3.5) a été considéré par Zhu (2011), Davis et Liu (2016) ainsi Christou et Fokianos (2014) qui ont donné pour $p = q = 1$ la condition de stationnarité stricte suivante

$$\alpha_{01}^2 \left(1 + \frac{1}{r_0}\right) + 2\alpha_{01}\beta_{01} + \beta_{01}^2 < 1,$$

avec un moment d'ordre deux fini. La formulation de Zhu (2011) est en effet,

$$X_t \mid \mathcal{F}_{t-1} \sim \mathcal{NB}\left(r_0, \frac{1}{1 + \mu_t}\right), \quad (1.3.6)$$

²proposée par Cameron et Trivedi (1986, 2013) dans le contexte de la régression statique à valeurs entières.

où $r_0 \in \mathbb{N}^*$ est restreint à être un entier positif et μ_t satisfaisant (1.3.1b). Toutefois, cette dernière peut être écrite sous la forme (1.3.5) en prenant $\lambda_t = \frac{\mu_t}{r_0}$.

Afin de mieux comprendre la dynamique de ce dernier, (1.3.6), considérons pour $p = q = 1$ le changement du bruit $\varepsilon_t = X_t - r_0\lambda_t$ où l'équation (1.3.1b) s'écrit explicitement comme suit

$$X_t - \varepsilon_t = r_0w_0 + r_0\alpha_{01}X_{t-1} + r_0\beta_{01}\lambda_{t-1},$$

où le terme d'erreur $\{\varepsilon_t\}$ est un bruit blanc. En effet :

$$\begin{aligned} E(\varepsilon_t) &= E(X_t - r_0\lambda_t) = E(E(X_t - r_0\lambda_t \mid \mathcal{F}_{t-1})) = 0. \\ \text{Var}(\varepsilon_t) &= \text{Var}(E(\varepsilon_t \mid \mathcal{F}_{t-1})) + E(\text{Var}(\varepsilon_t \mid \mathcal{F}_{t-1})) \\ &= E(\text{Var}(X_t \mid \mathcal{F}_{t-1})) = r_0E(\lambda_t(1 + \lambda_t)) \\ &= r_0\mu_\lambda + r_0E(\lambda_t^2) = \mu + \frac{\mu^2}{r_0} + r_0\text{Var}(\lambda_t), \end{aligned}$$

qui est indépendante de t puisque $\{X_t\}$ est stationnaire.

$$\text{Cov}(\varepsilon_t, \varepsilon_{t+h}) = E(\varepsilon_t\varepsilon_{t+h}) = E(\varepsilon_t E(\varepsilon_{t+h} \mid \mathcal{F}_{t+h-1})) = 0, \forall h > 0.$$

Ainsi, sous l'hypothèse que $r_0\alpha_{01} + \beta_{01} < 1$, les moments du processus $NB_2\text{-INGARCH}(1, 1)$ sont tels que :

$$1. E(X_t) = E(E(X_t \mid \mathcal{F}_{t-1})) = r_0E(\lambda_t) \equiv \mu = r_0\mu_\lambda = \frac{r_0w_0}{1 - (r_0\alpha_{01} + \beta_{01})}.$$

$$2. \text{var}(X_t) = \text{var}(\varepsilon_t + r_0\lambda_t) = \mu + \frac{\mu^2}{r_0} + (r_0 + r_0^2)\text{var}(\lambda_t), \text{ où}$$

$$\text{var}(\lambda_t) = E(\lambda_t^2) - E^2(\lambda_t) = \frac{\alpha_{01}^2}{1 - ((r_0\alpha_{01} + \beta_{01})^2 + r_0b_1^2)} \left(\mu + \frac{\mu^2}{r_0} \right).$$

D'où :

$$\text{var}(X_t) = \frac{1 - (r_0\alpha_{01} + \beta_{01})^2 + r_0^2b_1^2}{1 - (r_0\alpha_{01} + \beta_{01})^2 - r_0\alpha_{01}^2} \left(\mu + \frac{\mu^2}{r_0} \right).$$

3. De plus, l'autocovariance pour ce modèle est telle que :

$$\gamma_X(h) = E(X_tX_{t-h}) - \mu^2,$$

or : $E(X_t X_{t-h}) = E(E(X_t X_{t-h} | \mathcal{F}_{t-1})) = (r_0 \alpha_{01} + \beta_{01}) \gamma_X(h-1) + \mu^2$.

D'où,

$$\begin{aligned} \gamma_X(h) &= (r_0 \alpha_{01} + \beta_{01}) \gamma_X(h-1) \\ &= (r_0 \alpha_{01} + \beta_{01})^{h-1} \gamma_X(1), \quad h \geq 2 \end{aligned}$$

De même pour λ , on obtient :

$$\begin{aligned} \gamma_\lambda(h) &= (r_0 \alpha_{01} + \beta_{01}) \gamma_\lambda(h-1) \\ &= (r_0 \alpha_{01} + \beta_{01})^h \gamma_\lambda(0), \quad h \geq 1 \end{aligned}$$

D'autre part³, on a :

$$\begin{aligned} \gamma_X(1) &= r_0 \alpha_{01} \gamma_X(0) + r_0^2 \beta_{01} \gamma_\lambda(0) \\ &= r_0^2 (\alpha_{01} + r_0 \alpha_{01} + \beta_{01}) \gamma_\lambda(0) + \alpha_{01} (r_0 \mu + \mu^2) \end{aligned} \quad (1.3.7)$$

et :

$$\gamma_\lambda(0) = \frac{\alpha_{01}}{r_0} \gamma_X(1) + \beta_{01} \gamma_\lambda(1) \quad (1.3.8)$$

Ainsi, la résolution de (1.3.7) et (1.3.8) donne :

$$\begin{aligned} \gamma_\lambda(0) &= \frac{\alpha_{01}^2 \left(\mu + \frac{\mu^2}{r_0} \right)}{1 - (r_0 \alpha_{01} + \beta_{01})^2 - r_0 \alpha_{01}^2} \\ \gamma_X(1) &= \frac{r_0 \alpha_{01} (1 - \beta_{01} (r_0 \alpha_{01} + \beta_{01}))}{1 - (r_0 \alpha_{01} + \beta_{01})^2 - r_0 \alpha_{01}^2} \left(\mu + \frac{\mu^2}{r_0} \right) \end{aligned}$$

Finalement :

$$\gamma_X(h) = (r_0 \alpha_{01} + \beta_{01})^{h-1} \frac{r_0 \alpha_{01} (1 - \beta_{01} (r_0 \alpha_{01} + \beta_{01}))}{1 - (r_0 \alpha_{01} + \beta_{01})^2 - r_0 \alpha_{01}^2} \left(\mu + \frac{\mu^2}{r_0} \right), \quad h \geq 1$$

³voir théorème 3 de Zhu, F. (2011)

1.3.3 Modèle INGARCH log-linéaire

Bien que les modèles précédents semblent offrir un cadre adéquat pour la modélisation de données de comptage dépendantes, il y a au moins deux inconvénients liés à ses applications. Par exemple en considérant le modèle *INGARCH* poissonien, du fait que $0 < \alpha_{01} + \beta_{01} < 1$, $cov(X_t, X_{t+h}) > 0$ ce qui signifie que ce modèle a été conçu principalement pour la modélisation des corrélations positives. Donc, il ne peut être utilisé lorsque la structure de dépendance admet des corrélations négatives. Une lacune supplémentaire de ce modèle est qu'il peut comprendre exclusivement des variables qui se traduisent par un terme de régression positif, car sinon la moyenne de ce processus devient négative. Pour s'attaquer à ces problèmes, Fokianos et Tjøstheim (2011) ont proposé un modèle dans lequel la distribution conditionnelle de X_t compte tenu de ses valeurs passées est poissonienne avec comme paramètre d'intensité $\lambda_t = e^{v_t}$, où

$$\begin{aligned} X_t \mid \mathcal{F}_{t-1} &\sim \mathcal{Poisson}(\lambda_t), \\ v_t &= w_0 + \beta_{01}v_{t-1} + \alpha_{01} \log(X_{t-1} + 1), t \geq 1, \end{aligned} \tag{1.3.9}$$

cela conduit aux modèles log-linéaires poissoniens du premier ordre, où les paramètres $w_0, \alpha_{01}, \beta_{01}$ appartenant à \mathbb{R} , v_0 et X_0 sont supposés être fixés. Fokianos et Tjøstheim (2011) ont montré que sous la condition de stabilité suivante

$$\begin{aligned} |\alpha_{01} + \beta_{01}| &< 1, \quad \text{si } \alpha_{01} > 0, \\ |\beta_{01}| |\alpha_{01} + \beta_{01}| &< 1, \quad \text{sinon.} \end{aligned} \tag{1.3.10}$$

le processus $\{X_t, t \in \mathbb{Z}\}$ donné par (1.3.9) est strictement stationnaire et ergodique. Afin de mieux motiver le choix de la fonction $\log(\cdot)$, considérons un modèle ressemblant à (1.3.9) mais avec X_{t-1} au lieu de $\log(X_{t-1} + 1)$. En d'autres termes, soit :

$$\begin{aligned} X_t \mid \mathcal{F}_{t-1} &\sim \mathcal{Poisson}(\lambda_t), \\ v_t &= w_0 + \beta_{01}v_{t-1} + \alpha_{01}X_{t-1}, t \geq 1, \end{aligned}$$

mais dans ce cas,

$$\lambda_t = \exp(w_0) \lambda_{t-1}^{\beta_{01}} \exp(\alpha_{01}X_{t-1}).$$

Ainsi, la stabilité de ce dernier est vérifiée uniquement dans le cas où $\alpha_{01} < 0$.

Notons que les modèles log-linéaires de séries chronologiques de comptage, ont été considérés par plusieurs auteurs, y compris Zeger et Qaqish (1988), Li (1994), MacDonald et Zucchini (1997), Brumback et al (2000), Kedem et Fokianos (2002), Davis et al (2003), Fokianos et Kedem (2004), Jung et al (2006), et que le modèle (1.3.9) est lié aux travaux de Zeger et Qaqish (1988), Li (1994), mais ces derniers n'ont abordé ni le problème de l'ergodicité ni celui de l'inférence asymptotique.

Ainsi, ces modèles permettent d'enrichir la classe des modèles autorégressifs en fournissant des structures alternatives permettant de produire des combinaisons qui sont à la fois positives et négatives. De plus, ils sont plus avantageux, car ils peuvent inclure des variables d'une façon très facile.

La question d'avoir une expression explicite de la fonction d'autocorrélation pour le modèle (1.3.9), similaire à celle obtenue pour le modèle (1.3.1b.), semble être un problème passionnant. En effet, la difficulté du calcul provient de la forme de la représentation log-linéaire, car contrairement au cas précédent, on ne peut pas obtenir une relation de récurrence pour la covariance. Notons ainsi que les modèles log-linéaires peuvent produire des corrélations positives et des corrélations négatives.

Afin d'examiner empiriquement le comportement du processus (1.3.9), on a eu recours à des simulations. La *figure 1.3.3* exhibe, pour différentes configurations des paramètres, le tracé de 200 observations générées à partir du modèle (1.3.9) suivies par leurs fonctions d'autocorrélation simples associées. v_0 est choisi égal à 1 (les valeurs de départ n'ont pas d'influence sur les résultats). En effet, ces observations sont prises après avoir généré 500 données et omis les 300 premières afin d'y assurer les conditions de stationnarité.

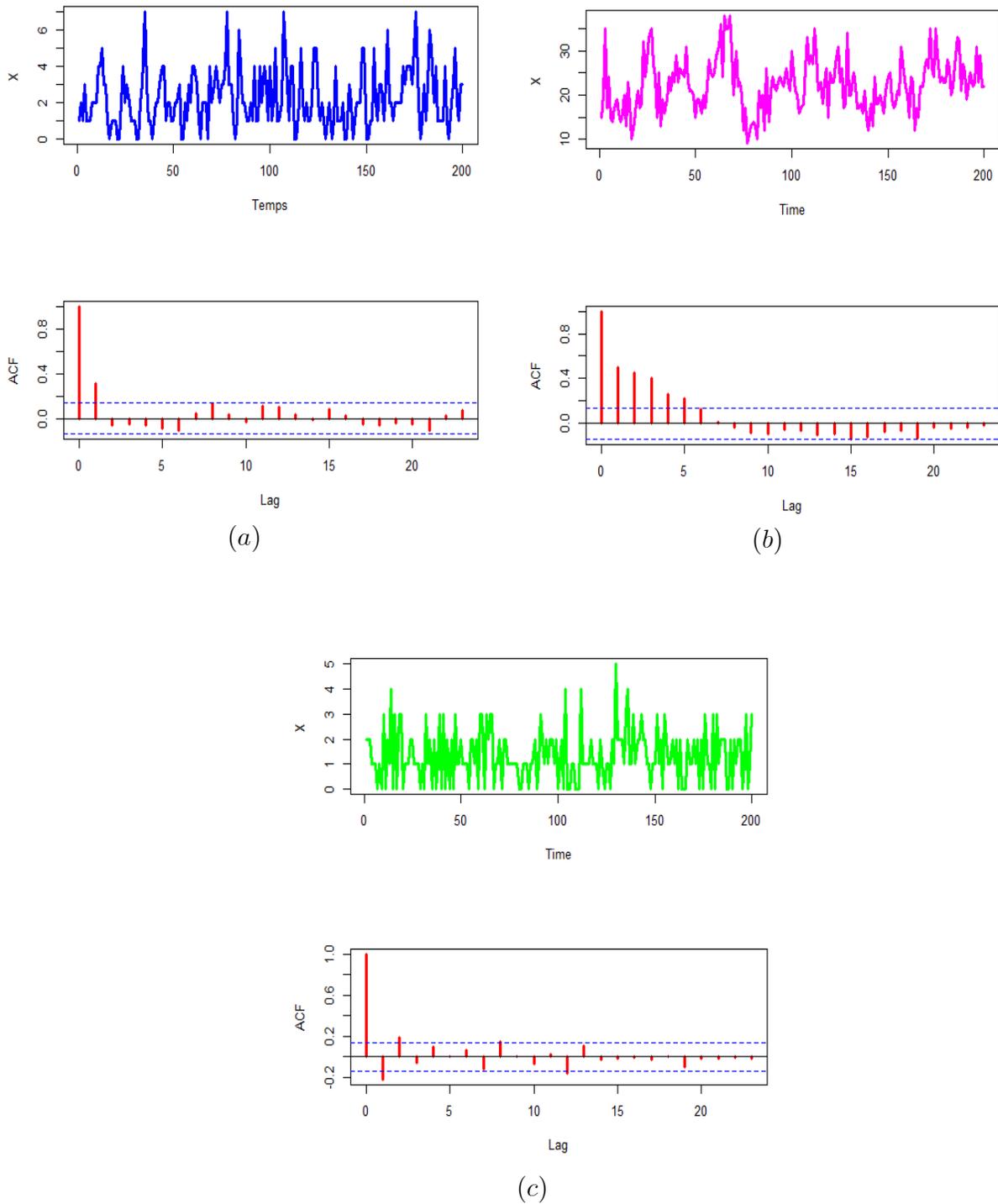


Figure 1.3.3 : 200 observations et ses fonctions d'autocorrélations simples à partir du modèle (1.3.9)

pour différentes valeurs des paramètres :

(a) : $w_0 = 0.5$, $\alpha_{01} = -0.5$, $\beta_{01} = 2/3$.

(b) : $w_0 = 0.5$, $\alpha_{01} = 1/3$, $\beta_{01} = 0.5$.

(c) : $w_0 = 0.5$, $\alpha_{01} = -1/4$, $\beta_{01} = -0.5$.

La *figure.1.3.3(a)* indique que lorsque α_{01} est positif et $|\alpha_{01} + \beta_{01}| < 1$, alors on obtient une série chronologique dont la fonction d'autocorrélation d'ordre 1 est positive mais la dépendance se dégrade pour des retards étendus. Toutefois, lorsque les paramètres α_{01} et β_{01} sont tous les deux positifs et leur somme correspondante est proche de 1, alors de la *Figure.1.3.3(b)* on observe une forte dépendance qui démontre que les valeurs du processus $\{X_t\}$ deviennent importantes. Lorsque α_{01} et β_{01} sont tous les deux négatifs alors la fonction d'autocorrélation d'ordre 1 est négative.

1.4 Modèles basés sur l'opérateur d'amincissement

Une classe de modèles, assez large, est celle qui est basée sur l'idée de l'amincissement. De tels modèles sont discutés par McKenzie (1985*a, b*; 1986; 1987; 1988*a, b*), Al-Osh et Alzaïd (1987; 1988; 1991), Du et Li (1991) et Al-Osh et Aly (1992). Dans cette section, nous introduisons ces modèles en considérant d'abord le cas le plus explicite, $INAR(1)$, ensuite le modèle $INAR(p)$ et à la fin nous présentons un cas un peu plus général, le modèle $GINAR(p)$. Notons que les résultats mentionnés ci-dessus, dans la section précédente, sont dans le cadre de certaines distributions marginales spécifiques tandis que les résultats de cette section ne requièrent aucune hypothèse distributionnelle explicite.

Généralement, les modèles autorégressifs à valeurs entières, $INAR$, sont basés essentiellement sur l'opérateur d'amincissement introduit par Steutel et Van Harn (1979). Ce dernier, dénoté par " \circ ", est connu sous le nom de "opérateur d'amincissement binomial". Il est défini comme suit

Définition 1.4.1 *Soit X une variable aléatoire à valeurs entières positives. Alors, l'opérateur d'amincissement binomial est tel que*

$$\alpha_0 \circ X = \sum_{j=1}^X Y_j,$$

où $\{Y_j, j \in \mathbb{N}\}$ est une suite de variables aléatoires de Bernoulli de paramètre $\alpha_0 \in (0, 1)$ indépendantes identiquement distribuées et indépendante de X . Cette suite est appelée série de comptage de $\alpha_0 \circ X$. Par conséquent, conditionnellement à X la variable aléatoire $\alpha_0 \circ X$ suit une loi binomiale de paramètres X et α_0 .

Maintenant, nous donnons quelques propriétés de base de l'opérateur d'amincissement binomial :

$$\begin{aligned} \cdot 0 \circ X &= 0, & \cdot E(\alpha_0 \circ X) &= \alpha_0 E(X), \\ \cdot 1 \circ X &= X, & \cdot \text{var}(\alpha_0 \circ X) &= \alpha_0^2 \text{var}(X) + \alpha_0(1 - \alpha_0) E(X), \\ \cdot \beta \circ (\alpha_0 \circ X) &= (\beta\alpha_0) \circ X, & \cdot \text{cov}(\alpha_0 \circ X, X) &= \alpha_0 \text{var}(X). \end{aligned}$$

Pour une preuve explicite de ces dernières et bien d'autres propriétés nous nous référons à Freeland (1998) et Da Silva (2005).

1.4.1 Modèle $INAR(1)$

Le modèle $INAR(1)$ a été proposé par McKenzie (1985) et étudié par la suite par Al-Osh et Alzaid (1987). Ainsi, ce dernier est défini comme suit

Définition 1.4.2 *Un processus du second ordre $\{X_t, t \in \mathbb{Z}\}$ vérifie un modèle autorégressif à valeurs entières d'ordre un, $INAR(1)$, s'il est de la forme*

$$\begin{cases} X_t = \alpha_0 \circ X_{t-1} + \varepsilon_t, & t \in \mathbb{Z}, \\ \alpha_0 \circ X_{t-1} = \sum_{j=1}^{X_{t-1}} Y_j \sim B(X_{t-1}, \alpha_0), \end{cases} \quad (1.4.1)$$

où le bruit (ε_t) est une suite de variables aléatoires *i.i.d* à valeurs dans \mathbb{N} , avec $E(\varepsilon_t) = \omega_0 > 0$ et $\text{var}(\varepsilon_t) = \sigma_0^2 > 0$, et (Y_j) est une suite de variables aléatoires indépendantes, identiquement distribuées suivant une loi de Bernoulli de paramètre $\alpha_0 \in (0, 1)$, indépendante de X_{t-1} et (ε_t) est indépendante de (Y_j) . Par conséquent, $(\alpha_0 \circ X_t | X_t)$ suit une loi binomiale de paramètres X_t et α_0 .

Sous l'hypothèse que $0 \leq \alpha_0 < 1$, les moments de ce modèle existent et sont tels que :

$$E(X_t) = \frac{\omega_0}{1 - \alpha_0}, \text{ et } \text{var}(X_t) = \frac{\sigma_0^2 + \omega_0 \alpha_0}{1 - \alpha_0^2}.$$

De plus, beaucoup de ses propriétés sont similaires à celles du processus $AR(1)$ standard, en particulier la fonction d'autocorrélation (ACF). Elle vaut :

$$\rho(h) = \text{corr}(X_t, X_{t-h}) = \alpha_0^h, \quad \forall h \in \mathbb{N}.$$

Ainsi, sous l'hypothèse que ε_t suit une loi de Poisson de paramètre λ , $E(\varepsilon_t) = \text{var}(\varepsilon_t) = \omega_0$, la loi stationnaire de la chaîne de Markov (X_t) est aussi de Poisson de paramètre $\frac{\omega_0}{1 - \alpha_0}$

et ce modèle sera appelé $PINAR(1)$. Pour d'autres propriétés concernant ce dernier, nous nous référons à Freeland (1998), Freeland et McCabe (2004, 2005) et Weib (2008).

Notons que la loi de Poisson n'est pas toujours appropriée à la modélisation et l'analyse des séries chronologiques à valeurs entières car la moyenne et la variance de la distribution de Poisson sont égales et que cette propriété n'est pas toujours vérifiée pour des données réelles. D'autres distributions marginales possibles pour le processus stationnaire $INAR(1)$ sont la loi binomiale négative (voir par exemple, Al-Osh et Aly (1992), Zhu et Joe (2006), Weib (2008)), la loi géométrique (voir par exemple, McKenzie (1986), Ristic et al. (2009)), et la loi de Poisson généralisée (voir Al-Osh et Alzaid (1993)).

1.4.2 Modèle $INAR(p)$

Une extension directe aux modèles autorégressifs à valeurs entières d'ordre p , $INAR(p)$, généralisant ainsi les résultats donnés par Al-Osh et Al-Zaïd (1987), $INAR(1)$, est donnée par la définition suivante

Définition 1.4.3 Soit $\{X_t\}_{t \in \mathbb{Z}}$, une suite de variables aléatoires à valeurs entières positives ; $\{\varepsilon_t\}_{t \in \mathbb{Z}}$, une suite de variables aléatoires i.i.d. à valeurs entières positives, avec $E(\varepsilon_t) = \omega > 0$ et $\text{var}(\varepsilon_t) = \sigma^2 > 0$, $p \in \mathbb{N}^*$ et $\{\alpha_j\}_{j \in \{1, \dots, p\}}$ une suite de constantes telles que $0 \leq \alpha_j < 1$, $\forall j$ et $(\alpha_p > 0)$. Alors, $\{X_t\}_{t \in \mathbb{Z}}$ est un processus $INAR(p)$ s'il est de la forme

$$X_t = \sum_{j=1}^p \alpha_j \circ X_{t-j} + \varepsilon_t, \quad \forall t \in \mathbb{Z}. \quad (1.4.2)$$

Ainsi, nous avons

$$\alpha_j \circ X_{t-j} = \sum_{k=1}^{X_{t-j}} Y_{jk},$$

où la série de comptage (Y_{jk}) est une suite de variables aléatoires indépendantes, identiquement distribuées selon une loi de Bernoulli de paramètres α_j , et indépendante de X_{t-j} . De plus, (ε_t) est indépendante de (Y_{jk}) . Notons qu'il existe deux structures particulières pour ce modèle ; l'une proposée par Al-Osh et Alzaïd (AA), (1990), et l'autre étudiée par Du &

Li (DL), (1991). Dans la première approche (AA), les auteurs supposent que la distribution conditionnelle du vecteur aléatoire $(\alpha_1 \circ X_t, \alpha_2 \circ X_t, \dots, \alpha_p \circ X_t)$ étant donné $X_t = x_t$ est multinomiale de paramètres $(\alpha_1, \dots, \alpha_p, X_t)$, et indépendante du processus passé, i.e, indépendante de X_{t-k} et de tous les amincissements de celui-ci, $\forall k > 0$. Par contre, dans la deuxième approche (DL), les auteurs supposent de plus que les séries de comptage sont mutuellement indépendantes.

Comme pour le modèle $INAR(1)$, les moments du processus $INAR(p)$ sont faciles à dériver sous la condition de stationnarité $\sum_{j=1}^p \alpha_j < 1$, qui est commune pour les deux approches.

Cependant, ces deux différentes formulations impliquent diverses structures du second ordre du processus : sous la première approche, (AA), le modèle $INAR(p)$ a la même structure de covariance que celle du modèle $ARMA(p, p - 1)$ tandis que sous la deuxième approche, (DL), cette structure du second ordre est similaire à celle du modèle $AR(p)$.

1.4.3 D'autres généralisations

Dans le but d'enrichir la classe des modèles autorégressifs permettant l'analyse des séries chronologiques à valeurs entières, Gautier & Latour (1994), Dion et al. (1995) et Latour (1998) ont considéré une version plus générale du modèle $INAR(p)$, noté $GINAR(p)$, basée sur l'opérateur généralisé de Steutel & Van Harn (1979).

Définition 1.4.4 (L'opérateur d'amincissement généralisé, noté $\alpha\star$). Soit X une variable aléatoire à valeurs entières positives ; Y , une variable aléatoire à valeurs entières positives de moyenne finie α et de variance finie λ et (Y_k) une suite de variables aléatoires indépendantes entre elles, indépendantes de X et distribuées selon la même loi que Y . Alors, l'opérateur $\alpha\star$ est défini par

$$\alpha\star X = \sum_{k=1}^X Y_k.$$

La suite (Y_k) est aussi une série de comptage. De plus, en supposant que $\beta\star$ est un autre opérateur d'amincissement généralisé de X , basé sur une série de comptage (δ_k) , alors les opérateurs $\alpha\star$ et $\beta\star$ sont indépendants si, et seulement si, les séries de comptage (Y_k) et

(δ_k) sont mutuellement indépendantes.

Définition 1.4.5 Soit $\{X_t\}_{t \in \mathbb{Z}}$, une suite de variables aléatoires à valeurs entières positives ; $\{\varepsilon_t\}_{t \in \mathbb{Z}}$, une suite de variables aléatoires i.i.d. à valeurs entières positives de moyenne finie μ_ε et de variance finie σ_ε^2 , $p \in \mathbb{N}^*$ et $\{\alpha_j\}_{j \in \{1, \dots, p\}}$ une suite de constantes telles que $0 \leq \alpha_j < 1$, $\forall j$ et $(\alpha_p > 0)$. Alors, $\{X_t\}_{t \in \mathbb{Z}}$ est un processus $GINAR(p)$ s'il est de la forme

$$X_t = \sum_{j=1}^p \alpha_j \star X_{t-j} + \varepsilon_t, \quad \forall t \in \mathbb{Z}. \quad (1.4.3)$$

Ici, toutes les séries de comptage $(Y_{jk})_{k \in \mathbb{N}^*}$ associées à $\alpha_j \star$ pour $j = 1, 2, \dots, p$ sont indépendantes entre elles et indépendantes de ε_t . Elles sont de moyenne finie α_j et de variance finie λ_j . De plus, les auteurs (*Gautier & Latour (1994)*) n'exigent pas qu'elles soient de type Bernoulli, ce qui distingue nettement leurs résultats de ceux de *Du & Li (1991)*.

Sous l'hypothèse $\sum_{j=1}^p \alpha_j < 1$, le processus $GINAR(p)$ est stationnaire et les moments du premier et du second ordre sont tels que

$$E(X_t) = \frac{\mu_\varepsilon}{1 - \sum_{j=1}^p \alpha_j} = \mu_X,$$

et

$$var(X_t) = \mu_X \sum_{j=1}^p \lambda_j + \sum_{j=1}^p \alpha_j \gamma(j) + \sigma_\varepsilon^2, \text{ où } \gamma(j) = cov(X_t, X_{t+j}), \forall t \in \mathbb{Z}.$$

Quant à la structure d'autocorrélation pour le processus $GINAR(p)$, elle est identique à celle d'un processus $AR(p)$ réel.

1.5 Approche *GLM* (generalized linear models) des modèles de séries chronologiques à valeurs entières

Au début du 20^{ème} siècle, l'analyse de données discrètes, telles que les données de comptage, s'est développée. Au panel des distributions disponibles pour la modélisation sont alors venues s'ajouter les distributions de Poisson ou binomiale.

La famille exponentielle permet de regrouper toutes ces lois et donne naissance à une nouvelle classe de modèles. Il s'agit des *modèles linéaires généralisés (GLM)*. Cette terminologie est introduite par Nelder et Wedderburn, 1972. Comme son nom l'indique, cette classe de modèles généralise les modèles linéaires classiques. Il s'agit d'une généralisation en termes de loi de probabilité d'une part, et en termes de lien à la linéarité. L'hypothèse sur la distribution associée à chaque modélisation est alors remplacée par une propriété de linéarité commune à tous les modèles, et par une relation moyenne conditionnelle-variance conditionnelle, i.e.

$$\text{Var}(X_t | \mathcal{F}_{t-1}) = l(E(X_t | \mathcal{F}_{t-1})), \quad (1.5.1)$$

où $l : (0, \infty) \rightarrow (0, \infty)$ est une fonction réelle positive. Dans la littérature des *modèles linéaires généralisés* (e.g. Nelder et Wedderburn, 1972 ; McCullagh et Nelder, 1989), une telle fonction de lien est aussi appelée *hypothèse de la variance nominale GLM* et est induite soit par la distribution conditionnelle du modèle quand elle est entièrement spécifiée, soit par la structure du modèle.

Par exemple, lorsque la distribution conditionnelle du modèle est poissonnienne de paramètre λ_t , ce qui se réduit à un cas particulier de l'autorégression poissonnienne générale proposée par Doukhan et al (2012), la fonction de lien *GLM* poissonnienne dans ce cas est donnée par la forme linéaire

$$l(x) = x.$$

Une fonction de lien linéaire plus générale

$$l(x) = \left(1 + \frac{1}{r_0}\right) x, \text{ pour certain } r_0 > 0,$$

est induite par la distribution conditionnelle binomiale négative₁, 1.3.4, vu que

$$\begin{aligned} E(X_t | \mathcal{F}_{t-1}) &= \lambda_t, \\ \text{var}(X_t | \mathcal{F}_{t-1}) &= \lambda_t \left(1 + \frac{1}{r_0}\right). \end{aligned}$$

De même, la fonction de lien impliquée par une distribution conditionnelle binomiale négative₂, c'est-à-dire $\mathcal{NB}\left(r_0, \frac{r_0}{r_0 + \lambda_t}\right)$, est donnée par

$$l(x) = x \left(1 + x \frac{1}{r_0}\right), \quad r_0 > 0. \quad (1.5.2)$$

Lorsque $r_0 = 1$, nous trouvons la fonction de lien correspondante à la distribution géométrique. D'autre part, une fonction de lien peut être trouvée même lorsque la distribution conditionnelle du modèle est mal spécifiée.

Dans la section 5.3.1, nous verrons que la fonction de lien *GLM* pour le modèle *INAR*(1) est toujours une fonction affine indépendamment de la distribution conditionnelle de ce modèle.

1.6 Généralisation : le modèle *INGAR* (INteger Gene-ralized AutoRegression)

Une classe assez générale de modèles de séries chronologiques à valeurs entières, comme proposé par Ahmad et Francq (2016), s'exprime à travers un processus observable à valeurs entières $\{X_t, t \in \mathbb{Z}\}$ défini sur un espace probabilisé (Ω, \mathcal{F}, P) avec une moyenne conditionnelle spécifiée comme suit

$$E(X_t | \mathcal{F}_{t-1}) = \lambda(X_{t-1}, X_{t-2}, \dots; \theta_0) := \lambda_t(\theta_0) := \lambda_t, \quad t \in \mathbb{Z}, \quad (1.6.1)$$

où $\mathcal{F}_t \subset \mathcal{F}$ est la σ -algèbre générée par $\{X_t, X_{t-1}, \dots\}$. Soit

$$\varepsilon_t := \varepsilon_t(\theta_0) = X_t - E(X_t | \mathcal{F}_{t-1}),$$

le modèle (1.6.1), qui est défini par la représentation moyenne conditionnelle (1.6.1), peut également être écrit comme l'équation de différence stochastique suivante (Ou sous forme d'innovation, voir. Grunwald et *al*, 2000)

$$X_t = \lambda(X_{t-1}, X_{t-2}, \dots; \theta_0) + \varepsilon_t, \quad t \in \mathbb{Z}. \quad (1.6.2)$$

L'équation (1.6.2), qui est pilotée par la différence $\{\mathcal{F}_t, t \in \mathbb{Z}\}$ -martingale $\{\varepsilon_t, t \in \mathbb{Z}\}$, semble être une autorégression généralisée infinie avec une solution à valeurs entières $\{X_t, t \in \mathbb{Z}\}$. Le terme "généralisé" se réfère à la forme générale de la fonction λ , qui peut être linéaire ou non linéaire. C'est pourquoi le modèle est appelé *INteger Generalized AutoRegression* (*INGAR*). En fait, le modèle *INGAR* (1.6.1)-(1.6.2) est assez général et englobe de nombreuses classes importantes de modèles de séries chronologiques à valeurs entières tels que le modèle (stable) *INGARCH* poissonien (Heinen, 2003 ; Ferland et al, 2006), l'autorégression poissonienne générale (Doukhan et al, 2012 ; Doukhan et Kengne, 2015 ; Kengne, 2015), le modèle *INGARCH* binomial négatif stable (Zhu, 2011 ; Christou et Fokianos, 2014 ; Davis et Liu, 2016 ; Diop et Kengne, 2017) et le modèle *INAR* (Al-Osh et Alzaid, 1987).

Notons que la généralisation du modèle *INGAR* (1.6.1) découle non seulement de la forme générale de la fonction $\lambda(\cdot)$ (voir aussi Doukhan et Wintenberger, 2008), mais aussi du fait qu'outre la moyenne conditionnelle, aucune autre spécification concernant la distribution conditionnelle du processus $\{X_t, t \in \mathbb{N}\}$ n'est nécessaire.

Dans le cadre de ce travail, nous nous sommes intéressés par l'estimation du paramètre inconnu θ_0 de la moyenne conditionnelle en utilisant la série X_1, X_2, \dots, X_n ($n \in \mathbb{N}^*$) générée à partir de (1.6.1). Au cas où la fonction de lien binomiale négative₂ comme (1.5.1)-(1.5.2) est spécifiée, nous nous sommes également intéressés par l'estimation du paramètre de dispersion r_0 . En fait, deux cas de (1.6.1) sont considérés

1^{er} cas : Seule la moyenne conditionnelle (1.6.1) est spécifiée de sorte que nous devons seulement estimer le paramètre θ_0 de la moyenne conditionnelle.

2^{ème} cas : L'équation (1.6.1) et la fonction de lien GLM binomiale négative₂ (1.5.1)-(1.5.2) sont toutes les deux spécifiées de sorte que nous devons estimer θ_0 et r_0 .

Une instance particulièrement importante du 2^{ème} cas apparaît lorsque la distribution conditionnelle du modèle est spécifiée entièrement comme étant binomiale négative₂, i.e.

$$X_t \mid \mathcal{F}_{t-1} \sim \mathcal{NB} \left(r_0, \frac{r_0}{r_0 + \lambda_t} \right),$$

où le modèle *INGARCH* binomial négatif₂ est un cas particulier (voir Davis et Liu, 2016 ; Zhu, 2011 ; Christou et Fokianos, 2014 et la section 1.3.2 ci-dessus). Pour notre estimation,

nous faisons l'hypothèse de régularité suivante

A0 Le processus $\{X_t, t \in \mathbb{Z}\}$ donné par (1.6.1) est strictement stationnaire et ergodique. Pour certaines classes particulières de (1.6.1) comme les modèles *INGARCH* et *INAR*, l'hypothèse **A0** peut être exprimée plus explicitement comme condition de stabilité sur θ_0 (voir Ahmad et Francq, 2016 et la section 1.3 ci-dessus). En outre, lorsque la distribution conditionnelle de (1.6.1) est poissonnienne, Doukhan et al (2012) ont fourni des conditions générales sur la fonction λ en (1.6.1) pour la stationnarité stricte et l'ergodicité du modèle. Maintenant, étant donné un paramètre générique $\theta \in \Theta$, la fonction moyenne conditionnelle donnée par

$$\lambda(X_{t-1}, X_{t-2}, \dots; \theta) := \lambda_t(\theta), \quad t \in \mathbb{N},$$

coïncide clairement avec la moyenne conditionnelle dans (1.6.1) lorsque $\theta = \theta_0$. Elle est inobservable en raison des valeurs non observables $X_0, X_{-1}, X_{-2}, \dots$. Pour toute valeur initiale fixe arbitraire $\tilde{X}_0, \tilde{X}_{-1}, \tilde{X}_{-2}, \dots$, soit

$$\tilde{\lambda}_t(\theta) = \lambda\left(X_{t-1}, X_{t-2}, \dots, X_1, \tilde{X}_0, \tilde{X}_{-1}, \dots; \theta\right), \quad t \in \mathbb{N}^*,$$

une approximation observable pour $\lambda_t(\theta)$. Cette dernière approximation sert à calculer divers types du *QMLE* de θ_0 que nous avons l'intention d'étudier ci-après.

Chapitre 2

Estimateur du quasi-maximum de vraisemblance poissonien du modèle INGAR

2.1 Introduction

Ce chapitre porte sur la consistance et la normalité asymptotique de *l'estimateur du quasi-maximum de vraisemblance poissonien* (P - $QMLE$), lequel est calculé comme si la distribution conditionnelle des observations était poissonnienne. Pour un exposé sur la méthode du QML et ses généralisations voir par exemple Aknouche (2013, 2014, 2015) et Aknouche et al (2018b). Nous allons synthétiser les principaux résultats que nous avons relevé auprès des travaux de Ahmad et Francq (2016) qui ont proposé un estimateur QML poissonien (P - $QMLE$) pour le paramètre θ_0 de la moyenne conditionnelle du modèle $INGAR$.

2.2 Estimateur du QML poissonien ($PQML$) du modèle $INGAR$

Ahmad et Francq (2016) ont proposé un estimateur QML poissonien (P - $QMLE$) pour le paramètre θ_0 de la moyenne conditionnelle du modèle $INGAR$ (1.6.1), qui est défini comme étant une solution mesurable au problème suivant

$$\hat{\theta}_P = \arg \max_{\theta \in \Theta} \left(\tilde{L}_{P,n}(\theta) \right), \quad (2.2.1a)$$

où

$$\tilde{L}_{P,n}(\theta) = \frac{1}{n} \sum_{t=1}^n \tilde{l}_t(\theta), \text{ et} \quad (2.2.1b)$$

$$\tilde{l}_t(\theta) = \left(-\tilde{\lambda}_t(\theta) + X_t \log \left(\tilde{\lambda}_t(\theta) \right) \right).$$

2.3 Propriétés asymptotiques de l'estimateur du PQML

Pour étudier la consistance du P - $QMLE$, $\hat{\theta}_P$, Ahmad et Francq (2016) ont posé les hypothèses suivantes :

A1 La fonction $\theta \mapsto \lambda_t(\theta)$ est continue *p.s.*; $\lambda_t(\theta) > c$ et $\tilde{\lambda}_t(\theta) > c$, *p.s.* pour un certain $c > 0$.

A2 $a_t \xrightarrow[t \rightarrow \infty]{p.s.} 0$ et $a_t X_t \xrightarrow[t \rightarrow \infty]{p.s.} 0$ où $a_t = \sup_{\theta \in \Theta} \left| \tilde{\lambda}_t(\theta) - \lambda_t(\theta) \right|$.

A3 $E(X_t^\delta) < \infty$ pour un certain $\delta > 1$.

A4 $\lambda_t(\theta) = \lambda_t(\theta_0)$ *p.s.* si et seulement si $\theta = \theta_0$.

A5 Θ est compact.

Les hypothèses **A1-A5** sont standards et peuvent être rendues plus explicites pour certains modèles particuliers de (1.6.1).

Théorème 2.3.1 (Ahmad et Francq, 2016) *Sous (1.6.1) et A0-A5,*

$$\hat{\theta}_P \xrightarrow{p.s.} \theta_0, \text{ lorsque } n \rightarrow \infty. \quad (2.3.1)$$

La normalité asymptotique du P - $QMLE$ a été étudiée par plusieurs auteurs (voir en particulier Christou et Fokianos (2014), etc.). Le cadre général de la moyenne conditionnelle de la forme (1.6.1) n'a cependant jamais été considéré. Fokianos et Tjøstheim (2012) ont considéré le MLE pour un modèle non linéaire poissonien de premier ordre. Ahmad et Francq (2016) ont donné des conditions pour la normalité asymptotique du P - $QMLE$ lorsque la vraie distribution conditionnelle n'est pas spécifiée. Ces conditions sont les suivantes.

A6 Les variables c_t , $c_t X_t$, $a_t d_t$ et $a_t d_t X_t$ sont d'ordre $O(t^{-\tau})$ p.s. pour un certain $\tau > 1/2$

$$\text{où } c_t = \sup_{\theta \in \Theta} \left\| \frac{\partial \left(\tilde{\lambda}_t(\theta) - \lambda_t(\theta) \right)}{\partial \theta} \right\| \text{ et } d_t = \sup_{\theta \in \Theta} \max \left(\left\| \frac{1}{\tilde{\lambda}_t(\theta)} \frac{\partial \tilde{\lambda}_t(\theta)}{\partial \theta} \right\|, \left\| \frac{1}{\lambda_t(\theta)} \frac{\partial \lambda_t(\theta)}{\partial \theta} \right\| \right).$$

A7 La vraie valeur θ_0 est à l'intérieur de Θ .

A8 La variance conditionnelle $v_t(\theta_0) := \text{Var}(X_t | \mathcal{F}_{t-1}) = E(X_t^2 | \mathcal{F}_{t-1}) - \lambda_t^2(\theta_0)$ est finie p.s.

A9 Les dérivées $\frac{\partial^2 \lambda_t(\theta)}{\partial \theta \partial \theta'}$ et $\frac{\partial^2 \tilde{\lambda}_t(\theta)}{\partial \theta \partial \theta'}$ existent et sont continues, les matrices

$$I_P = E \left(\frac{v_t(\theta_0)}{\lambda_t^2(\theta_0)} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right) \text{ et } J_P = E \left(\frac{1}{\lambda_t(\theta_0)} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right),$$

sont finies et J_P est non singulière.

A10 Il existe un voisinage $V(\theta_0)$ de θ_0 tel que $E \left(\sup_{\theta \in V(\theta_0)} \left\| \frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'} \right\| \right) < \infty$.

Une propriété importante du P - $QMLE$ est sa robustesse vis-à-vis d'une mauvaise spécification de la vraie distribution conditionnelle du modèle (1.6.1).

Théorème 2.3.2 (Ahmad et Francq, 2016) Sous (1.6.1) et **A0-A10**,

$$\sqrt{n} \left(\hat{\theta}_P - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, J_P^{-1} I_P J_P^{-1} \right), \quad (2.3.2)$$

Ahmad et Francq (2016) ont montré que sous les hypothèses du théorème 2.3.2, la variance asymptotique du P - $QMLE$ peut être estimée par $\hat{J}^{-1} \hat{I} \hat{J}^{-1}$ avec

$$\hat{I}_P = \frac{1}{n} \sum_{t=1}^n \left(\frac{X_t}{\tilde{\lambda}_t(\hat{\theta}_p)} - 1 \right)^2 \frac{\partial \tilde{\lambda}_t(\hat{\theta}_p)}{\partial \theta \partial \theta'} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_p)}{\partial \theta \partial \theta'}. \quad (2.3.3)$$

$$\hat{J}_P = \frac{1}{n} \sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_p)} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_p)}{\partial \theta \partial \theta'} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_p)}{\partial \theta \partial \theta'}. \quad (2.3.4)$$

2.4 Applications à des modèles de séries chronologiques à valeurs entières spécifiques

Cette section a pour objet de clarifier les résultats cités précédemment à propos de la distribution asymptotique de l'estimateur quasi-maximum de vraisemblance poissonien à travers deux exemples de modèles de séries chronologiques à valeurs entières ; à savoir le modèle *INGARCH*(1,1) et le modèle *INAR*(1). Et ce, tout en gardant comme distribution initiale la loi de Poisson ensuite nous utiliserons une distribution binomiale négative $_K$. Pour cette dernière, nous explicitons les résultats en s'appuyant sur la forme introduite par Cameron et Trividi (1986, 2013) ($K = 1, K = 2$).

Comme nous l'avons déjà précisé précédemment, les hypothèses \mathbf{A}_0 - \mathbf{A}_{10} proposées par Ahmad et Francq 2016 peuvent être écrite sous format explicite pour des cas particuliers du modèle *INGAR*. Alors, dans le cadre du *PQMLE* et pour clarifier d'abord l'hypothèse \mathbf{A}_0 , nous donnons l'espace de stationnarité correspondant à chaque cas particulier. Ensuite, nous établissons les propriétés de la consistance et de la normalité asymptotique de l'estimateur pour ces cas particuliers de modèles. Enfin, nous donnons leurs matrices de variance asymptotiques ce qui clarifie ainsi l'hypothèse \mathbf{A}_9 . Pour le reste des hypothèses, nous considérons comme exemple le cas du modèle Poisson *INGARCH* : les hypothèses \mathbf{A}_1 et \mathbf{A}_4 découlent de la forme linéaire de λ_t dans (1.3.1) et puisque (1.3.2) entraîne l'existence des moments de tout ordre (Ferland et al.,2006), alors les hypothèses $\mathbf{A}_3, \mathbf{A}_8$ et \mathbf{A}_9 sont évidemment vérifiées (voir aussi Francq et Zakoian,2010). Un argument similaire montre que les hypothèses \mathbf{A}_2 et \mathbf{A}_6 sont vérifiées.

2.4.1 Application au modèle *INGARCH* poissonien

Pour le processus *INGARCH* (1,1) poissonien, nous définissons l'espace de stationnarité par

$$\Theta_1 = \{\theta \in \mathbb{N}^2 \mid \alpha_{01} + \beta_{01} < 1\}.$$

Si $\theta_0 \in \Theta_1$, alors ce processus est strictement stationnaire et ergodique ce qui veut dire que l'hypothèse \mathbf{A}_0 est bien vérifiée. Ainsi, nous pouvons énoncer la proposition suivante

Proposition 2.4.1 (Ahmad et Francq, 2016) *Soit $\theta_0 \in \Theta$ un sous ensemble compact de Θ_1 et X une solution stationnaire de (1.3.1). Alors*

1. *l'EQMV poissonien est fortement consistant $\hat{\theta}_P \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$.*
2. *$\sqrt{n}(\hat{\theta}_P - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, J_P^{-1})$, où*

$$J_P = E \left(\frac{1}{\lambda_t(\theta_0)} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

Dans ce cas, la matrice de variance asymptotique estimée est telle que

$$V(\hat{\theta}_P) \simeq n^{-1} \hat{J}_P^{-1} = \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_P)} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta'} \right)^{-1}.$$

2.4.2 Application au modèle INGARCH binomial négatif_K

En considérant le cas du modèle NB_1 -INGARCH(1,1), nous avons le résultat suivant

Proposition 2.4.2 (Ahmad et Francq, 2016) Soit $\theta_0 \in \Theta$ un sous ensemble compact de Θ_1 et X une solution stationnaire de (1.3.4). Alors, sous **A0-A10**

1. l'EQMV poissonien est fortement consistant $\widehat{\theta}_P \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$.
2. $\sqrt{n}(\widehat{\theta}_P - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, J_P^{-1} I_P J_P^{-1})$, où les matrices I_P et J_P sont telles que

$$I_P = E \left(\left(1 + \frac{1}{r_0} \right) \frac{1}{\lambda_t(\theta_0)} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right),$$

$$J_P = E \left(\frac{1}{\lambda_t(\theta_0)} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

En considérant le cas d'une distribution binomiale négative₂, l'espace de stationnarité est tel que

$$\Theta_2 = \left\{ \theta \in \mathbb{N}^2 \mid \alpha_0^2 \left(1 + \frac{1}{r_0} \right) + 2\alpha_0\beta_0 + \beta_0^2 < 1 \right\}.$$

Si $\theta_0 \in \Theta_2$, alors la stationnarité du processus NB_2 -INGARCH(1,1) est assurée et les hypothèses **A1-A10** sont trivialement vérifiées.

Proposition 2.4.3 (Ahmad et Francq, 2016) Soit $\theta_0 \in \Theta$ un sous ensemble compact de Θ_2 et X une solution stationnaire de (1.3.5). Alors,

1. l'EQMV poissonien est fortement consistant $\widehat{\theta}_P \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$.
2. $\sqrt{n}(\widehat{\theta}_P - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, J_P^{-1} I_P J_P^{-1})$, où les matrices I_P et J_P sont telles que

$$I_P = E \left(\left(\frac{1}{\lambda_t(\theta_0)} + \frac{1}{r_0} \right) \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right),$$

$$J_P = E \left(\frac{1}{\lambda_t(\theta_0)} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

Et la matrice de variance estimée est comme suit

$$\begin{aligned}
 V\left(\widehat{\theta}_P, r_0\right) &\simeq n^{-1} \widehat{J}_P^{-1} \widehat{I}_P \widehat{J}_P^{-1} \\
 &= \left(\sum_{t=1}^n \frac{1}{\widetilde{\lambda}_t\left(\widehat{\theta}_P\right)} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta'} \right)^{-1} \left(\sum_{t=1}^n \frac{\left(r_0 + \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)\right)}{r_0 \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta'} \right) \\
 &\quad \times \left(\sum_{t=1}^n \frac{1}{\widetilde{\lambda}_t\left(\widehat{\theta}_P\right)} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta'} \right)^{-1}.
 \end{aligned}$$

Remarque 2.4.1

Pour le P -QMLE avec INGARCH géométrique, les matrices I_P et J_P sont estimées par

$$\begin{aligned}
 I_P &\simeq n^{-1} \sum_{t=1}^n \frac{\left(1 + \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)\right)}{\lambda_t\left(\widehat{\theta}_P\right)} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta'}, \\
 J_P &\simeq n^{-1} \sum_{t=1}^n \frac{1}{\widetilde{\lambda}_t\left(\widehat{\theta}_P\right)} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta'}.
 \end{aligned}$$

Et la matrice de variance estimée est comme suit

$$\begin{aligned}
 V\left(\widehat{\theta}_G\right) &\simeq n^{-1} \widehat{J}_P^{-1} \widehat{I}_P \widehat{J}_P^{-1} \\
 &= \left(\sum_{t=1}^n \frac{1}{\widetilde{\lambda}_t\left(\widehat{\theta}_P\right)} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta'} \right)^{-1} \\
 &\quad \times \left(\sum_{t=1}^n \frac{\left(1 + \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)\right)}{\widetilde{\lambda}_t\left(\widehat{\theta}_P\right)} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta'} \right) \left(\sum_{t=1}^n \frac{1}{\widetilde{\lambda}_t\left(\widehat{\theta}_P\right)} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t\left(\widehat{\theta}_P\right)}{\partial \theta'} \right)^{-1}.
 \end{aligned}$$

2.4.3 Application au modèle $INAR(1)$

Dans le cas du modèle $INAR(1)$, nous savons que

$$v_t(\theta_0) = \alpha_0(1 - \alpha_0)X_{t-1} + \sigma_0^2,$$

et l'espace de stationnarité est donné par

$$\Theta_3 = \{\theta \in \mathbb{N} \mid \alpha_{01} < 1\}.$$

Si $\theta_0 \in \Theta_3$, alors ce processus est strictement stationnaire et ergodique ce qui veut dire que l'hypothèse \mathbf{A}_0 est bien vérifiée. Ainsi, nous pouvons énoncer la proposition suivante

Proposition 2.4.4 (Ahmad et Francq, 2016) *Soit $\theta_0 \in \Theta$ un sous ensemble compact de Θ_3 et X une solution stationnaire de (1.4.1). Alors*

1. *l'EQMV poissonien est fortement consistant $\widehat{\theta}_P \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$.*

2. *$\sqrt{n}(\widehat{\theta}_P - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, J_P^{-1}I_P J_P^{-1})$, où les matrices I_P et J_P sont telles que*

$$I_P = E \left(\frac{\alpha_0(1 - \alpha_0)X_{t-1} + \sigma_0^2}{\lambda_t^2(\theta_0)} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right),$$

$$J_P = E \left(\frac{1}{\lambda_t(\theta_0)} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

Dans ce cas, la matrice de variance asymptotique estimée est telle que

$$\begin{aligned} V(\widehat{\theta}_P) &\simeq n^{-1} \widehat{J}_P^{-1} \widehat{I}_P \widehat{J}_P^{-1} = \left(\sum_{t=1}^n \frac{1}{\widetilde{\lambda}_t(\widehat{\theta}_P)} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_P)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_P)}{\partial \theta'} \right)^{-1} \\ &\quad \times \left(\sum_{t=1}^n \frac{\widehat{\alpha}_P(1 - \widehat{\alpha}_P)X_{t-1} + \widehat{\sigma}_P^2}{\lambda_t^2(\widehat{\theta}_P)} \frac{\partial \lambda_t(\widehat{\theta}_P)}{\partial \theta} \frac{\partial \lambda_t(\widehat{\theta}_P)}{\partial \theta'} \right) \\ &\quad \times \left(\sum_{t=1}^n \frac{1}{\widetilde{\lambda}_t(\widehat{\theta}_P)} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_P)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_P)}{\partial \theta'} \right)^{-1}. \end{aligned}$$

INAR avec innovation poissonnienne

Lorsque $\varepsilon_t \sim \mathcal{P}(w_0)$, ce qui veut dire $\sigma_0^2 = w_0$. Alors,

$$I_P = E \left(\frac{\alpha_0 (1 - \alpha_0) X_{t-1} + w_0 \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'}}{\lambda_t^2(\theta_0)} \right).$$

Ainsi, la matrice de variance estimée est donnée par

$$\begin{aligned} V(\hat{\theta}_P) &\simeq \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_P)} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta'} \right)^{-1} \\ &\quad \times \left(\sum_{t=1}^n \frac{\hat{\alpha}_P (1 - \hat{\alpha}_P) X_{t-1} + \hat{w}_P \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta'}}{\tilde{\lambda}_t^2(\hat{\theta}_P)} \right) \\ &\quad \times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_P)} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta'} \right)^{-1}. \end{aligned}$$

INAR avec innovation géométrique

Lorsque $\varepsilon_t \sim \mathcal{G}\left(\frac{1}{w_0}\right)$, ce qui veut dire $\sigma_0^2 = w_0(w_0 - 1)$. Donc,

$$I_P = E \left(\frac{\alpha_0 (1 - \alpha_0) X_{t-1} + w_0 (w_0 - 1) \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'}}{\lambda_t^2(\theta_0)} \right).$$

Ainsi, une estimation consistante de la matrice de variance asymptotique est donnée par

$$\begin{aligned} V(\hat{\theta}_P) &\simeq \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_P)} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta'} \right)^{-1} \\ &\quad \times \left(\sum_{t=1}^n \frac{\hat{\alpha}_P (1 - \hat{\alpha}_P) X_{t-1} + \hat{w}_P (\hat{w}_P - 1) \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta'}}{\lambda_t^2(\hat{\theta}_P)} \right) \\ &\quad \times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_P)} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta'} \right)^{-1}. \end{aligned}$$

INAR avec innovation binomiale négative₂

Lorsque $\varepsilon_t \sim \mathcal{NB}\left(r, \frac{r}{r+w_0}\right)$, ce qui veut dire $\sigma_0^2 = \frac{1}{r}w_0(r+w_0)$. Donc,

$$I_P = E \left(\frac{\alpha_0(1-\alpha_0)X_{t-1} + \frac{1}{r}w_0(r+w_0)}{\lambda_t^2(\theta_0)} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

Alors, la matrice de variance théorique estimée est telle que

$$\begin{aligned} V(\hat{\theta}_P, r) &\simeq \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_P)} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta'} \right)^{-1} \\ &\times \left(\sum_{t=1}^n \frac{\hat{\alpha}_P(1-\hat{\alpha}_P)X_{t-1} + \frac{1}{r}\hat{w}_0(r+\hat{w}_0)}{\tilde{\lambda}_t^2(\hat{\theta}_P)} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta'} \right) \\ &\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_P)} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_P)}{\partial \theta'} \right)^{-1}. \end{aligned}$$

Chapitre 3

Estimateur du quasi-maximum de vraisemblance géométrique (*GQML*) du modèle *INGAR*

3.1 Introduction

Dans ce chapitre nous nous intéressons principalement à étudier, sous un jeu d'hypothèses assez faibles, la consistance et la normalité asymptotique de notre estimateur proposé "estimateur du quasi-maximum de vraisemblance géométrique (*GQMLE*)". Ce dernier appartient à la même classe de modèles étudiée par Ahmad et Francq (2016), calculé comme si la distribution conditionnelle du processus observable était géométrique. Bien entendu, l'estimateur ne serait asymptotiquement efficace que si la distribution conditionnelle des observations était géométrique. Cependant, comme cette loi est surdispersée l'estimateur du quasi-maximum de vraisemblance calculé à partir d'elle peut avoir un bon niveau d'efficacité asymptotique relative en présence de séries réelles surdispersées.

3.2 Estimateur du *QML* géométrique (*GQML*) du modèle *INGAR*

Dans cette section, nous nous intéressons à l'estimation de θ_0 partant d'une série X_1, X_2, \dots, X_n générée selon la forme (1.6.1) qui est soumise à l'hypothèse **A0**. En se basant sur la loi géomé-

trique de paramètre $\frac{1}{1 + \tilde{\lambda}_t(\theta)}$, la fonction de vraisemblance géométrique du modèle (1.6.1)

est donnée après transformation logarithmique par

$$\tilde{L}_{G,n}(\theta) = \frac{1}{n} \sum_{t=1}^n \tilde{l}_t(\theta) \text{ avec } \tilde{l}_t(\theta) = \log\left(\frac{1}{1+\tilde{\lambda}_t(\theta)}\right) + X_t \log\left(\frac{\tilde{\lambda}_t(\theta)}{1+\tilde{\lambda}_t(\theta)}\right), \theta \in \Theta, \quad (3.2.1)$$

et un estimateur du quasi-maximum de vraisemblance géométrique (*GQMLE*) de θ_0 est le maximum de $\tilde{L}_{G,n}(\theta)$ sur Θ , c.-à-d. une solution mesurable du problème

$$\hat{\theta}_G = \arg \max_{\theta \in \Theta} \left(\tilde{L}_{G,n}(\theta) \right)$$

pour un certain espace paramétrique Θ . Par ailleurs, du fait que la fonction de vraisemblance donnée par (3.2.1) appartient à la famille de lois exponentielle, elle coïncide avec la fonction du quasi-vraisemblance de Wedderburn (1974) qui est basée sur l'hypothèse de modèle linéaire généralisé (*GLM*) suivante

$$\text{Var}(X_t | \mathcal{F}_{t-1}) = E(X_t | \mathcal{F}_{t-1})(1 + E(X_t | \mathcal{F}_{t-1})) = \lambda_t(1 + \lambda_t). \quad (3.2.2)$$

3.3 Propriétés asymptotiques de l'estimateur du *GQML*

Pour étudier la consistance du *GQMLE*, nous gardons les mêmes hypothèses **A1-A5**, celles posées par Ahmad et Francq (2016) pour la consistance de leur *PQMLE*.

Théorème 3.3.1 (Aknouche et Bendjeddou, 2017a) *Sous (1.6.1) et A0-A5, nous avons*

$$\hat{\theta}_G \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0.$$

Le résultat précédent montre que, comme le *PQMLE*, le *GQMLE* est également robuste à une mauvaise spécification de la vraie loi conditionnelle du modèle, pourvu que la moyenne conditionnelle soit bien spécifiée, ce qui n'est pas surprenant puisque la vraisemblance géométrique appartient à la famille de lois exponentielle (cf. Gourieroux et al, 1984). Pour établir la normalité asymptotique du *GQMLE*, posons $l_t(\theta)$ et $L_{G,n}(\theta)$ les variables aléatoires obtenues à partir de $\tilde{l}_t(\theta)$ et $\tilde{L}_{G,n}(\theta)$ respectivement en remplaçant $\tilde{\lambda}_t(\theta)$ par $\lambda_t(\theta)$ et

considérons les mêmes hypothèses de régularité supplémentaires précédentes (proposées par Ahmad et Francq (2016) **A6-A10**) mais nous proposons **A'6** et **A'9** au lieu de **A6** et **A9** :

A'6 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 $O(t^{-\tau})$ p.s. pour un

certain $\tau > 1/2$ où $b_t = \sup_{\theta \in \Theta} \left| \tilde{\lambda}_t^2(\theta) - \lambda_t^2(\theta) \right|$, $c_t = \sup_{\theta \in \Theta} \left\| \frac{\partial \left(\tilde{\lambda}_t(\theta) - \lambda_t(\theta) \right)}{\partial \theta} \right\|$ et

$$d_t = \sup_{\theta \in \Theta} \max \left(\left\| \frac{1}{\tilde{\lambda}_t(\theta) (1 + \tilde{\lambda}_t(\theta))} \frac{\partial \tilde{\lambda}_t(\theta)}{\partial \theta} \right\|, \left\| \frac{1}{\lambda_t(\theta) (1 + \lambda_t(\theta))} \frac{\partial \lambda_t(\theta)}{\partial \theta} \right\| \right).$$

A'9 Les dérivées $\frac{\partial^2 \lambda_t(\theta)}{\partial \theta \partial \theta'}$ et $\frac{\partial^2 \tilde{\lambda}_t(\theta)}{\partial \theta \partial \theta'}$ existent et sont continues, 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.

Théorème 3.3.2 (Aknouche et Bendjedou, 2017a) Sous (1.6.1), **A0-A5**, **A'6**, **A7-A8**, **A'9** et **A10**, nous avons

$$\sqrt{n} \left(\hat{\theta}_G - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, J_G^{-1} I_G J_G^{-1} \right). \quad (3.3.1)$$

Un estimateur consistant de la variance asymptotique $J_G^{-1} I_G J_G^{-1}$ de $\hat{\theta}_G$ est $\hat{J}_G^{-1} \hat{I}_G \hat{J}_G^{-1}$ avec

$$\hat{I}_G \simeq \frac{1}{n} \sum_{t=1}^n \left(\frac{X_t - \tilde{\lambda}_t(\hat{\theta}_G)}{\tilde{\lambda}_t(\hat{\theta}_G) (1 + \tilde{\lambda}_t(\hat{\theta}_G))} \right)^2 \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'}$$

et

$$\hat{J}_G \simeq \frac{1}{n} \sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_G) (1 + \tilde{\lambda}_t(\hat{\theta}_G))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'}.$$

3.4 Applications à des modèles à valeurs entières spécifiques

Dans le même but que la section 2.4, nous explicitons sous format de propositions les résultats liés à l'inférence statistique des estimateurs à travers des exemples et ce en reprenant les deux modèles particuliers déjà sélectionnés.

3.4.1 Application au modèle INGARCH poissonien

Proposition 3.4.1 (*Aknouche et Bendjeddou, 2017*) Soit $\theta_0 \in \Theta$ un sous ensemble compact de Θ_1 et X une solution stationnaire de (1.3.1). Alors

1. l'EQMV géométrique est fortement consistant $\hat{\theta}_G \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$.
2. $\sqrt{n} \left(\hat{\theta}_G - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, J_G^{-1} I_G J_G^{-1} \right)$, où les matrices I_G et J_G sont telles que

$$I_G = E \left(\frac{1}{\lambda_t(\theta_0) (1 + \lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right),$$

$$J_G = E \left(\frac{1}{\lambda_t(\theta_0) (1 + \lambda_t(\theta_0))} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

Dans ce cas, une estimation consistante de la matrice de variance asymptotique est telle que

$$V \left(\hat{\theta}_G \right) \simeq n^{-1} \hat{J}_P^{-1} \hat{I}_P \hat{J}_P^{-1} = \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t \left(\hat{\theta}_G \right) \left(1 + \tilde{\lambda}_t \left(\hat{\theta}_G \right) \right)} \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta} \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta'} \right)^{-1}$$

$$\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t \left(\hat{\theta}_G \right) \left(1 + \tilde{\lambda}_t \left(\hat{\theta}_G \right) \right)^2} \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta} \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta'} \right)$$

$$\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t \left(\hat{\theta}_G \right) \left(1 + \tilde{\lambda}_t \left(\hat{\theta}_G \right) \right)} \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta} \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta'} \right)^{-1}.$$

3.4.2 Application au modèle INGARCH binomial négatif₂

Proposition 3.4.2 (Aknouche et Bendjeddou, 2017) Soit $\theta_0 \in \Theta$ un sous ensemble compact de Θ_2 et X une solution stationnaire de (1.3.5). Alors,

1. l'EQMV géométrique est fortement consistant $\hat{\theta}_G \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$.
2. $\sqrt{n}(\hat{\theta}_G - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, J_G^{-1} I_G J_G^{-1})$, où les matrices I_G et J_G sont telles que

$$I_G = E \left(\frac{r_0 + \lambda_t(\theta_0)}{\lambda_t(\theta_0)(1 + \lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right),$$

$$J_G = E \left(\frac{1}{\lambda_t(\theta_0)(1 + \lambda_t(\theta_0))} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

Dans ce cas, la matrice de variance asymptotique estimée est telle que

$$\begin{aligned} V(\hat{\theta}_G, r_0) &\simeq n^{-1} \hat{J}_P^{-1} \hat{I}_P \hat{J}_P^{-1} = \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_G)(1 + \tilde{\lambda}_t(\hat{\theta}_G))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'} \right)^{-1} \\ &\times \left(\sum_{t=1}^n \frac{r_0 + \lambda_t(\theta_0)}{\lambda_t(\theta_0)(1 + \lambda_t(\theta_0))^2} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'} \right) \\ &\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_G)(1 + \tilde{\lambda}_t(\hat{\theta}_G))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'} \right)^{-1}. \end{aligned}$$

Remarque 3.4.1 Dans le cas d'un modèle INGARCH géométrique, la matrice I est donnée par l'expression suivante

$$I_G = E \left(\frac{1}{\lambda_t(\theta_0)(1 + \lambda_t(\theta_0))} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

Dans ce cas, la matrice de variance théorique estimée est donnée par

$$V(\hat{\theta}_G) \simeq n^{-1} \hat{J}_P^{-1} = \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_G)(1 + \tilde{\lambda}_t(\hat{\theta}_G))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'} \right)^{-1}.$$

3.4.3 Application au modèle INAR(1)

Proposition 3.4.1 (Aknouche et Bendjedou, 2017) Soit $\theta_0 \in \Theta$ un sous ensemble compact de Θ_3 et X une solution stationnaire de (1.4.1). Alors

1. l'EQMV géométrique est fortement consistant $\hat{\theta}_G \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$.
2. $\sqrt{n} \left(\hat{\theta}_G - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, J_G^{-1} I_G J_G^{-1} \right)$, où les matrices I_G et J_G sont telles que

$$I_G = E \left(\frac{\alpha_0 (1 - \alpha_0) X_{t-1} + \sigma_0^2 \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'}}{\lambda_t^2(\theta_0) (1 + \lambda_t(\theta_0))^2} \right) \text{ et}$$

$$J_G = E \left(\frac{1}{\lambda_t(\theta_0) (1 + \lambda_t(\theta_0))} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

La matrice de variance asymptotique estimée est telle que

$$\begin{aligned} V \left(\hat{\theta}_G \right) &\simeq n^{-1} \hat{J}_G^{-1} \hat{I}_G \hat{J}_G^{-1} \\ &= \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t \left(\hat{\theta}_G \right) \left(1 + \tilde{\lambda}_t \left(\hat{\theta}_G \right) \right)} \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta} \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta'} \right)^{-1} \\ &\quad \times \left(\sum_{t=1}^n \frac{\hat{\alpha}_G (1 - \hat{\alpha}_G) X_{t-1} + \hat{\sigma}_G^2 \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta} \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta'}}{\tilde{\lambda}_t^2 \left(\hat{\theta}_G \right) \left(1 + \tilde{\lambda}_t \left(\hat{\theta}_G \right) \right)^2} \right) \\ &\quad \times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t \left(\hat{\theta}_G \right) \left(1 + \tilde{\lambda}_t \left(\hat{\theta}_G \right) \right)} \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta} \frac{\partial \tilde{\lambda}_t \left(\hat{\theta}_G \right)}{\partial \theta'} \right)^{-1}. \end{aligned}$$

INAR avec innovation poissonnienne

Lorsque $\varepsilon_t \sim \mathcal{P}(w_0)$, ce qui veut dire $\sigma_0^2 = w_0$. Donc,

$$I_G = E \left(\frac{\alpha_0 (1 - \alpha_0) X_{t-1} + w_0 \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'}}{\lambda_t^2(\theta_0) (1 + \lambda_t(\theta_0))^2} \right).$$

Ainsi, la matrice de variance estimée est comme suit

$$\begin{aligned}
 V(\hat{\theta}_G) &\simeq \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_G)(1+\tilde{\lambda}_t(\hat{\theta}_G))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'} \right)^{-1} \\
 &\times \left(\frac{\hat{\alpha}_G(1-\hat{\alpha}_G)X_{t-1} + \hat{w}_G}{\tilde{\lambda}_t(\hat{\theta}_G)(1+\tilde{\lambda}_t(\hat{\theta}_G))^2} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'} \right) \\
 &\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_G)(1+\tilde{\lambda}_t(\hat{\theta}_G))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'} \right)^{-1}.
 \end{aligned}$$

INAR avec innovation géométrique

Lorsque $\varepsilon_t \sim \mathcal{G}\left(\frac{1}{w_0}\right)$, ce qui veut dire $\sigma_0^2 = w_0(w_0 - 1)$. Donc,

$$I_G = E \left(\frac{\alpha_0(1-\alpha_0)X_{t-1} + w_0(w_0-1)}{\lambda_t^2(\theta_0)(1+\lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

Ainsi, une estimation consistante de la matrice de variance asymptotique est donnée par

$$\begin{aligned}
 V(\hat{\theta}_G) &\simeq \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_G)(1+\tilde{\lambda}_t(\hat{\theta}_G))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'} \right)^{-1} \\
 &\times \left(\sum_{t=1}^n \frac{\hat{\alpha}_G(1-\hat{\alpha}_G)X_{t-1} + \hat{w}_G(\hat{w}_G-1)}{\tilde{\lambda}_t^2(\hat{\theta}_G)(1+\tilde{\lambda}_t(\hat{\theta}_G))^2} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'} \right) \\
 &\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_G)(1+\tilde{\lambda}_t(\hat{\theta}_G))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_G)}{\partial \theta'} \right)^{-1}.
 \end{aligned}$$

INAR avec innovation binomiale négative₂

Lorsque $\varepsilon_t \sim \mathcal{NB}\left(r, \frac{r}{r+w_0}\right)$, ce qui veut dire $\sigma_0^2 = \frac{1}{r}w_0(r+w_0)$. Donc,

$$I_r = E \left(\frac{\alpha_0(1-\alpha_0)X_{t-1} + \frac{1}{r}w_0(r+w_0)}{\lambda_t^2(\theta_0)(1+\lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

Alors, la matrice de variance théorique estimée est telle que

$$\begin{aligned} V(\hat{\theta}_r, r) &\simeq \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_r)(1+\tilde{\lambda}_t(\hat{\theta}_r))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right)^{-1} \\ &\times \left(\sum_{t=1}^n \frac{\hat{\alpha}_r(1-\hat{\alpha}_r)X_{t-1} + \frac{1}{r}\hat{w}_r(r+\hat{w}_r)}{\tilde{\lambda}_t^2(\hat{\theta}_r)(1+\tilde{\lambda}_t(\hat{\theta}_r))^2} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right) \\ &\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_r)(1+\tilde{\lambda}_t(\hat{\theta}_r))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right)^{-1}. \end{aligned}$$

3.5 Preuves des principaux résultats

Preuve du Théorème 3.3.1 En suivant l'approche de Wald, la preuve du Théorème 3.3.1 repose sur les deux lemmes suivants.

Lemme A.1 Sous **A1-A2**, $\sup_{\theta \in \Theta} \left| L_{G,n}(\theta) - \tilde{L}_{G,n}(\theta) \right| \xrightarrow[n \rightarrow \infty]{p.s.} 0$.

Preuve En utilisant l'inégalité $\log(x) \leq x - 1$, le fait que $\tilde{\lambda}_t(\theta) > 0$, les hypothèses **A1-A2** et puis le lemme de Césaro il vient que

$$\begin{aligned} \sup_{\theta \in \Theta} \left| L_{G,n}(\theta) - \tilde{L}_{G,n}(\theta) \right| &= \frac{1}{n} \sup_{\theta \in \Theta} \left| \sum_{t=1}^n \left(\log \left(\frac{\lambda_t(\theta) - \tilde{\lambda}_t(\theta)}{1 + \tilde{\lambda}_t(\theta)} + 1 \right) + X_t \log \left(\frac{\tilde{\lambda}_t(\theta) - \lambda_t(\theta)}{\lambda_t(\theta)(1 + \tilde{\lambda}_t(\theta))} + 1 \right) \right) \right| \\ &\leq \frac{1}{n} \sum_{t=1}^n \left(a_t + X_t a_t \frac{1}{c} \right) \xrightarrow[n \rightarrow \infty]{p.s.} 0. \quad \blacksquare \end{aligned}$$

Lemme A.2 Sous **A0-A1** et **A3-A4** : *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 D'abord, sous **A1** 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 admettent des moments finis de tout ordre et par les inégalités de Jensen et de Hölder ainsi que **A3** il s'ensuit que

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

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

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

où l'on peut observer que par (A.1) et (A.2), $E(l_1(\theta) - l_1(\theta_0)) \in [-\infty, 0]$ 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 **A4**, si et seulement si $\theta = \theta_0$. ■

Pour compléter la preuve du théorème soit $V_k(\bar{\theta})$ ($k \in \mathbb{N}^*$, $\bar{\theta} \in \Theta$) la boule ouverte de centre $\bar{\theta}$ et de rayon $1/k$. Puisque $\sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} l_t(\theta)$ est une fonction mesurable des termes de $\{X_t, t \in \mathbb{Z}\}$, processus strictement stationnaire et ergodique sous **A0**, il s'ensuit que $\left\{ \sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} l_t(\theta), t \in \mathbb{Z} \right\}$ est également strictement stationnaire et ergodique où l'on peut constater par le Lemme A.2 que $E \left(\sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} l_t(\theta) \right) \in [-\infty, +\infty[$. Ainsi, au vu du Lemme A.1 et le théorème ergodique (Billingsley, 2008) il vient que

$$\limsup_{n \rightarrow \infty} \sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} \tilde{L}_{G,n}(\theta) = \limsup_{n \rightarrow \infty} \sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} L_{G,n}(\theta) \leq E \left(\sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} l_1(\theta) \right),$$

et par le théorème de Beppo-Levi, $E \left(\sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} l_1(\theta) \right)$ décroît vers $E(l_1(\bar{\theta}))$ lorsque $k \rightarrow \infty$. Ainsi, nous avons montré qu'à l'égard de Lemme A.1 et Lemme A.2, il existe pour tout $\bar{\theta} \neq \theta_0$ un voisinage $V(\bar{\theta})$ tel que

$$\limsup_{n \rightarrow \infty} \sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} \tilde{L}_{G,n}(\theta) < \limsup_{n \rightarrow \infty} \tilde{L}_{G,n}(\theta_0) = \limsup_{n \rightarrow \infty} L_{G,n}(\theta_0) = E(l_1(\theta_0)).$$

Enfin, par des arguments standards la preuve du théorème s'achève en exploitant l'hypothèse **A5** de compacité de Θ . ■

Preuve du Théorème 3.3.2 D'abord par **A6**

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

Ensuite, comme par **A7** et Théorème 3.3.1 $\hat{\theta}_G$ ne peut être sur le bord de Θ pour n assez grand il s'ensuit par la formule de Taylor que pour un certain θ^* entre $\hat{\theta}_G$ et θ_0

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

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

Lemme A.3 Sous **A8-A'9**, $\sqrt{n} \frac{\partial L_{G,n}(\theta_0)}{\partial \theta} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, I_G)$.

Preuve Il est clair que $\left\{ \sqrt{n} \frac{\partial L_{G,n}(\theta_0)}{\partial \theta}, t \in \mathbb{Z} \right\}$ est une martingale par rapport à $\{\mathcal{F}_t, t \in \mathbb{Z}\}$ avec $\sqrt{n} \frac{\partial L_{G,n}(\theta_0)}{\partial \theta} = \sum_{t=1}^n \frac{1}{\sqrt{n}} \frac{\partial l_t(\theta_0)}{\partial \theta}$, $\frac{\partial l_t(\theta_0)}{\partial \theta} = \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{X_t - \lambda_t(\theta_0)}{\lambda_t(\theta_0)(1 + \lambda_t(\theta_0))}$ et par **A8-A'9** nous avons $E \left(\frac{\partial l_t(\theta_0)}{\partial \theta} \frac{\partial l_t(\theta_0)}{\partial \theta'} \right) = I_G$. Ainsi le lemme s'ensuit par le théorème central limite des martingales (ex. Billingsley, 2008). ■

Lemme A.4 Sous **A8-A10**, $\frac{\partial^2 L_{G,n}(\theta^*)}{\partial \theta \partial \theta'} \xrightarrow[n \rightarrow \infty]{p.s.} J_G$.

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 $V(\theta_0)$ défini par **A10**. Supposons également que n soit assez grand pour que θ^* soit dans $V_k(\theta_0)$. Donc par stationnarité et ergodicité de

$\sup_{\theta \in V_k(\theta_0)} \left| \frac{\partial^2 l_t(\theta)}{\partial \theta_i \partial \theta_j} - E \left(\frac{\partial^2 l_t(\theta_0)}{\partial \theta_i \partial \theta_j} \right) \right|$ nous avons

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

Au vu de **A10**, le théorème de convergence dominée de Lebesgue entraîne que

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

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

Chapitre 4

Estimation du quasi-maximum de vraisemblance binomial négatif (*NB-QML*) du modèle *INGAR*

4.1 Introduction

Ce chapitre considère deux estimateurs du quasi-maximum de vraisemblance binomial négatif du modèle *INGAR* (1.6.1) tout en admettant la réalisation X_1, \dots, X_n de ce modèle. Pour décrire ces estimateurs, considérons le 2^{ème} cas du modèle (1.6.1)-(1.5.2) avec les paramètres inconnus θ_0 et r_0 . Pour tout générique $\theta \in \Theta$ et $r > 0$, la fonction (log) vraisemblance binomiale négative, $\tilde{L}_{NB}(\theta, r)$, basée sur la distribution conditionnelle binomiale négative₂, $\mathcal{NB}\left(r, \frac{r}{r + \tilde{\lambda}_t(\theta)}\right)$, est donnée par

$$\tilde{L}_{NB}(\theta, r) = \frac{1}{n} \sum_{t=1}^n \tilde{l}_t(\theta, r), \quad (4.1.1)$$

$$\text{avec } \tilde{l}_t(\theta, r) = r \log\left(\frac{r}{r + \tilde{\lambda}_t(\theta)}\right) + X_t \log\left(\frac{\tilde{\lambda}_t(\theta)}{r + \tilde{\lambda}_t(\theta)}\right) + \frac{\Gamma(X_t + r)}{X_t! \Gamma(r)}.$$

Une estimation quasi-maximum de vraisemblance binomiale négative (*NB-QMLE*) de (θ_0, r_0) est une maximisation de $\tilde{L}_{NB}(\theta, r)$ sur $\theta \in \Theta$ et $r > 0$. Notons cependant que $\tilde{l}_t(\theta, r)$ donnée par (4.1.1) n'appartient pas à la famille exponentielle linéaire au sens de Gourieroux et al

(1984a). Ainsi, toute maximisation de (4.1.1) peut être inconsistante devant une mauvaise-spécification de la vraie distribution conditionnelle du modèle (1.6.1), ce qui constitue une sérieuse limitation. Au lieu de maximiser directement (4.1.1) et de reprendre la composante estimée correspondante à θ_0 , on peut considérer une approche à quatre étapes qui est plus robuste à la spécification incorrecte de la vraie distribution conditionnelle et qui consiste à :

i) Fixer r dans (4.1.1) arbitrairement à tout nombre positif connu, disons $r^* > 0$ et estimer θ_0 tout en maximisant (4.1.1) par rapport à θ , ce qui donne la première étape du *QMLE* $\hat{\theta}_{r^*}$.

ii) Estimer r_0 en se basant sur la relation *GLM* (1.5.1)-(1.5.2) en utilisant une estimation par moindres carrés pondérés \hat{r}_1 tout en remplaçant θ_0 par son *QMLE*, $\hat{\theta}_{r^*}$, obtenu en *i)*.

iii) Réestimer θ_0 en maximisant une variation de (4.1.1) obtenue en remplaçant r par son estimation \hat{r}_1 obtenue en *ii)*, ce qui donne $\hat{\theta}_{\hat{r}_1}$.

iv) Réestimer r_0 en utilisant la même méthode des moindres carrés pondérés utilisée en *ii)* mais en remplaçant θ_0 par $\hat{\theta}_{\hat{r}_1}$ obtenu en *iii)*.

Pour une approche similaire dans le contexte de la régression statique de comptage, voir Gourieroux et al (1984a, 1984b) et Wooldridge (1997, 2002). Dans les première et troisième étapes ci-dessus, la maximisation de (4.1.1) est effectuée par rapport à θ tout en gardant r fixé. Ainsi, le dernier terme de (4.1.1) peut être omis et (4.1.1) est remplacée simplement par "*Vraisemblance binomiale négative en Profil*"

$$\tilde{L}_{n,r}(\theta) = \frac{1}{n} \sum_{t=1}^n \tilde{l}_{t,r}(\theta) \quad \text{avec} \quad \tilde{l}_{t,r}(\theta) = r \log \left(\frac{r}{r + \tilde{\lambda}_t(\theta)} \right) + X_t \log \left(\frac{\tilde{\lambda}_t(\theta)}{r + \tilde{\lambda}_t(\theta)} \right). \quad (4.1.2)$$

Il convient de noter que $\tilde{l}_{t,r}(\theta)$ dans (4.1.2) appartient à la famille exponentielle linéaire. Par conséquent, toute maximisation de (4.1.2) par rapport à θ serait robuste en regard d'une spécification erronée de la distribution conditionnelle, chaque fois que l'on spécifie correctement la moyenne conditionnelle comme (1.6.1). Il s'avère que pour tout $r > 0$ fixé, $\tilde{L}_{n,r}(\theta)$ est la fonction du quasi-vraisemblance de Wedderburn (Wedderburn, 1974) basée sur la fonction de lien *GLM* binomiale négative (1.5.1)-(1.5.2) (avec r au lieu de r_0).

D'autre part, en considérant le 1^{er} cas du modèle (1.6.1) où seule la moyenne conditionnelle est spécifiée, alors seulement θ_0 doit être estimé et r dans (4.1.1) peut être fixé à toute valeur réelle positive. Ainsi, la maximisation de (4.1.1) ne sera faite qu'à l'égard de θ , ce qui revient à maximiser (4.1.2). En résumé, pour les deux cas **1** et **2**, nous devons maximiser la vraisemblance (ou Quasi-) en profil (4.1.2) par rapport à θ .

Dans le reste de ce chapitre, nous étudierons le comportement asymptotique de deux estimateurs de type *QML* qui maximisent (4.1.2) sur $\theta \in \Theta$. La section 4.3 examine la consistance et la normalité asymptotique d'une maximisation de (4.1.2) pour un r arbitrairement fixé et strictement positif. L'estimateur résultant sera appelé estimateur du quasi-maximum de vraisemblance binomial négatif (*pNB-QMLE*) en profil (ou marginal). Dans la section 4.6, la consistance et la normalité asymptotique de l'approche à quatre étapes (voir *i*) -*iv*) ci-dessus) sont établies en considérant la fonction de lien *GLM* (1.5.1)-(1.5.2) pour un $r_0 > 0$ inconnu.

4.2 Estimateur du *QML* binomial négatif en profil (*pNB-QML*)

Considérons le 1^{er} cas du modèle *INGAR* où seule (1.6.1) est requise. Un estimateur du quasi-maximum de vraisemblance binomial négatif en profil (*pNB-QMLE*) de θ_0 est une solution mesurable du problème suivant

$$\hat{\theta}_r = \arg \max_{\theta \in \Theta} \left(\tilde{L}_{n,r}(\theta) \right), \quad (4.2.1)$$

pour un certain Θ et un certain $r > 0$ connu fixé, où $\tilde{L}_{n,r}(\theta)$ est donnée par (4.1.2). Lorsque $r = 1$, $\hat{\theta}_1$ se réduit au *QMLE* géométrique (*G-QMLE*) étudié par Aknouche et Bendjeddou (2017). Le choix de $(\tilde{X}_0, \tilde{X}_{-1}, \dots)$ n'a pas d'importance asymptotique, mais peut influencer la précision de l'estimation en échantillons finis. En général, on suppose que $\tilde{X}_0 = x, \tilde{X}_{-1} = x, \dots$ Avec x dépendant de la fonction λ ou des observations (voir Ahmad et Francq, 2016).

4.3 Propriétés asymptotiques du $QMLE$ binomial négatif en profil

Pour étudier la consistance du $pNB-QMLE$, $\hat{\theta}_r$, nous avons besoin comme précédemment des hypothèses **A1-A5**, posées par Ahmad et Francq (2016) pour la consistance de leur $PQMLE$.

Théorème 4.3.1 (Aknouche, Bendjedou et Touche, 2018) *Sous (1.6.1) et **A0-A5**, nous avons*

$$\hat{\theta}_r \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0, \quad \text{pour tout } r > 0. \quad (4.3.1)$$

Le résultat précédent montre que, comme le $PQMLE$, le $pNB-QMLE$ est également robuste à une mauvaise spécification de la vraie loi conditionnelle du modèle, pourvu que la moyenne conditionnelle soit bien spécifiée, ce qui n'est pas surprenant puisque la log-vraisemblance binomiale négative en profil appartient à la famille de lois exponentielle (cf. Gouriéroux et al, 1984).

Examinons maintenant la normalité asymptotique du $pNB-QMLE$ tout en gardant la même définition de $\tilde{l}_{t,r}(\theta)$ donnée par (4.1.2) pour $l_{t,r}(\theta)$ avec $\lambda_t(\theta)$ au lieu de $\tilde{\lambda}_t(\theta)$, et soit

$$L_{n,r}(\theta) = \frac{1}{n} \sum_{t=1}^n l_{t,r}(\theta).$$

Et considérons les mêmes hypothèses de régularité supplémentaires précédentes (proposées par Ahmad et Francq (2016)) aux modifications suivantes près : nous remplaçons **A6** et **A9** par **A"6** et **A"9** :

A"6 *Les variables $c_t, c_t X_t, a_t d_t, a_t d_t X_t$ et $b_t d_t X_t$ sont d'ordre $O(t^{-\tau})$ p.s. pour un certain*

$\tau > 1/2$ où $b_t = \sup_{\theta \in \Theta} \left| \tilde{\lambda}_t^2(\theta) - \lambda_t^2(\theta) \right|$,

$$c_t = \sup_{\theta \in \Theta} \left\| \frac{\partial \left(\tilde{\lambda}_t(\theta) - \lambda_t(\theta) \right)}{\partial \theta} \right\|, \text{ et}$$

$$d_t = \sup_{\theta \in \Theta} \max \left(\left\| \frac{1}{\tilde{\lambda}_t(\theta) (r + \tilde{\lambda}_t(\theta))} \frac{\partial \tilde{\lambda}_t(\theta)}{\partial \theta} \right\|, \left\| \frac{1}{\lambda_t(\theta) (r + \lambda_t(\theta))} \frac{\partial \lambda_t(\theta)}{\partial \theta} \right\| \right).$$

A"9 Les dérivées $\frac{\partial^2 \lambda_t(\theta)}{\partial \theta \partial \theta'}$ et $\frac{\partial^2 \tilde{\lambda}_t(\theta)}{\partial \theta \partial \theta'}$ existent et sont continues, les matrices

$$I_r = E \left(\frac{v_t(\theta_0)}{\lambda_t^2(\theta_0) (r + \lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right) \text{ et}$$

$$J_r = E \left(\frac{1}{\lambda_t(\theta_0) (r + \lambda_t(\theta_0))} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right),$$

sont finies et J_r est non singulière pour tout $r > 0$.

Maintenant nous avons le résultat de normalité asymptotique suivant.

Théorème 4.3.2 (Aknouche, Bendjeddou et Touche, 2018) Sous (1.6.1), **A0-A5**, **A"6**, **A7-A8**, **A"9** et **A10** nous avons

$$\sqrt{n} \left(\hat{\theta}_r - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, J_r^{-1} I_r J_r^{-1} \right) \quad \text{pour tout } r > 0. \quad (4.3.2)$$

Remarque 4.3.1 : Voici quelques remarques :

1°. Lorsque la distribution conditionnelle du processus de génération de données (1.6.1) est \mathcal{NB}_2 avec les paramètres r_0 et $\frac{r_0}{r_0 + \lambda_t}$, i.e. $X_t \mid \mathcal{F}_{t-1} \sim \mathcal{NB} \left(r_0, \frac{r_0}{r_0 + \lambda_t} \right)$, alors (4.3.2) est vérifiée avec

$$I_r = \frac{1}{r_0} E \left(\frac{r_0 + \lambda_t(\theta_0)}{\lambda_t(\theta_0) (r + \lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

En particulier, lorsque r dans (4.1.2)-(4.2.1) coïncide avec la "vraie" valeur r_0 dans (1.5.1)-(1.5.2), alors $\hat{\theta}_{r_0}$ devient l'estimateur du maximum de vraisemblance (*MLE*) qui est alors

asymptotiquement efficace avec

$$I_{r_0} = \frac{1}{r_0} J_{r_0}.$$

Par conséquent, (4.3.2) devient

$$\sqrt{n} \left(\hat{\theta}_{r_0} - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, \frac{1}{r_0} J_{r_0}^{-1} \right). \quad (4.3.3)$$

2°. Un résultat plus faible, qui ne nécessite pas une spécification entière de la distribution conditionnelle, est que sous la fonction de lien *GLM* binomiale négative₂, qui généralise (1.5.1)-(1.5.2), suivante

$$\text{Var}(X_t | \mathcal{F}_{t-1}) = \delta^2 E(X_t | \mathcal{F}_{t-1}) \left(1 + \frac{1}{r_0} E(X_t | \mathcal{F}_{t-1}) \right) \text{ pour certain } \delta^2 > 0, r_0 > 0, \quad (4.3.4)$$

$\hat{\theta}_{r_0}$ est asymptotiquement efficace dans la classe de tous les estimateurs *QML* appartenant à la famille exponentielle (voir Gourieroux et al (1984a, 1984b) et Wooldridge (1997) dans le contexte de l'inférence du *QML* pour des modèles de régression statique à valeurs entières). Dans ce cas, nous avons

$$\sqrt{n} \left(\hat{\theta}_{r_0} - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, \delta^2 J_{r_0}^{-1} \right). \quad (4.3.5)$$

Toutefois, notons que r_0 est généralement inconnu et que (4.3.3) et (4.3.5) ne restent vrai que si r_0 est constamment estimé sous (4.3.4) comme on le verra dans la section 4.5.

Soulignons à présent la question importante consistant à estimer la variance asymptotique du *pNB-QMLE*. Comme pour Ahmad et Francq (2016), une estimation consistante de la variance asymptotique $J_r^{-1} I_r J_r^{-1}$ du *pNB-QMLE*, $\hat{\theta}_r$, est $\hat{J}_r^{-1} \hat{I}_r \hat{J}_r^{-1}$ avec

$$\hat{I}_r = \frac{1}{n} \sum_{t=1}^n \left(\frac{X_t - \tilde{\lambda}_t(\hat{\theta}_r)}{\tilde{\lambda}_t(\hat{\theta}_r) (r + \tilde{\lambda}_t(\hat{\theta}_r))} \right)^2 \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r) \partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta \partial \theta'}. \quad (4.3.6)$$

$$\hat{J}_r = \frac{1}{n} \sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_r) (r + \tilde{\lambda}_t(\hat{\theta}_r))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r) \partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta \partial \theta'}. \quad (4.3.7)$$

4.4 Application à des modèles spécifiques

Dans ce qui suit, nous étudions la consistance et la normalité asymptotique des estimateurs à travers des exemples. Ces applications sont considérées pour les deux modèles $INGARCH(1, 1)$ et $INAR(1)$.

Notons que dans le cadre du modèle $pNB-QMLE$, les hypothèses \mathbf{A}_0 - \mathbf{A}_{10} se simplifient comme dans le cas du modèle Poisson $INGARCH$ pour $p = q = 1$. Une différence notable intervient dans l'existence du moment d'ordre quatre, i.e, on doit supposer que $E(X_t^4) < \infty^1$. Lorsque p et q sont quelconques, une condition d'existence des moments plus générale semble être difficile à trouver.

4.4.1 Application au modèle $INGARCH$ poissonien

Proposition 4.4.1 (*Aknouche, Bendjeddou et Touche, 2018*) Soit $\theta_0 \in \Theta$ un sous ensemble compact de Θ_1 et X une solution stationnaire de (1.3.1). Alors

1. Le $pNB-QMLE$ est fortement consistant $\hat{\theta}_r \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$ pour tout $r > 0$.

2. $\sqrt{n} \left(\hat{\theta}_r - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, J_r^{-1} I_r J_r^{-1})$ pour tout $r > 0$, où les matrices I_r et J_r sont

telles que

$$I_r = E \left(\frac{1}{\lambda_t(\theta_0) (r + \lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right) \text{ et}$$

$$J_r = E \left(\frac{1}{\lambda_t(\theta_0) (r + \lambda_t(\theta_0))} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

¹ donnée par Ahmad et Francq (2016).

Dans ce cas, la matrice de variance asymptotique estimée est telle que

$$\begin{aligned}
 V\left(\hat{\theta}_r\right) &\simeq n^{-1} \hat{J}_r^{-1} \hat{I}_r \hat{J}_r^{-1} = \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t\left(\hat{\theta}_r\right)\left(r+\tilde{\lambda}_t\left(\hat{\theta}_r\right)\right)} \frac{\partial \tilde{\lambda}_t\left(\hat{\theta}_r\right)}{\partial \theta} \frac{\partial \tilde{\lambda}_t\left(\hat{\theta}_r\right)}{\partial \theta'} \right)^{-1} \\
 &\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t\left(\hat{\theta}_r\right)\left(1+\tilde{\lambda}_t\left(\hat{\theta}_r\right)\right)^2} \frac{\partial \tilde{\lambda}_t\left(\hat{\theta}_r\right)}{\partial \theta} \frac{\partial \tilde{\lambda}_t\left(\hat{\theta}_r\right)}{\partial \theta'} \right) \\
 &\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t\left(\hat{\theta}_r\right)\left(r+\tilde{\lambda}_t\left(\hat{\theta}_r\right)\right)} \frac{\partial \tilde{\lambda}_t\left(\hat{\theta}_r\right)}{\partial \theta} \frac{\partial \tilde{\lambda}_t\left(\hat{\theta}_r\right)}{\partial \theta'} \right)^{-1}.
 \end{aligned}$$

4.4.2 Application au modèle INGARCH binomial négatif₂

Proposition 4.4.2 (*Aknouche, Bendjeddou et Touche, 2018*) Soit $\theta_0 \in \Theta$ un sous ensemble compact de Θ_2 et X une solution stationnaire de (1.3.5). Alors,

1. Le pNB -QMLE est fortement consistant $\hat{\theta}_r \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$ pour tout $r > 0$.

2. $\sqrt{n}\left(\hat{\theta}_r - \theta_0\right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N\left(0, J_r^{-1} I_r J_r^{-1}\right)$ pour tout $r > 0$, où les matrices I_r et J_r sont

telles que

$$I_r = \frac{1}{r_0} E\left(\frac{1}{\lambda_t\left(\theta_0\right)\left(r+\lambda_t\left(\theta_0\right)\right)} \frac{\partial \lambda_t\left(\theta_0\right)}{\partial \theta} \frac{\partial \lambda_t\left(\theta_0\right)}{\partial \theta'}\right) = \frac{1}{r_0} J_r.$$

Dans ce cas, une estimation consistante de la matrice de variance asymptotique est telle que

$$V\left(\hat{\theta}_r, r_0\right) \simeq n^{-1} \frac{1}{r_0} \hat{J}_r^{-1} = \frac{1}{r_0} \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t\left(\hat{\theta}_r\right)\left(r+\tilde{\lambda}_t\left(\hat{\theta}_r\right)\right)} \frac{\partial \tilde{\lambda}_t\left(\hat{\theta}_r\right)}{\partial \theta} \frac{\partial \tilde{\lambda}_t\left(\hat{\theta}_r\right)}{\partial \theta'} \right)^{-1}.$$

4.4.3 Application au modèle $INAR(1)$

Proposition 4.4.3 (Aknouche, Bendjeddou et Touche, 2018) Soit $\theta_0 \in \Theta$ un sous ensemble compact de Θ_3 et X une solution stationnaire de (1.4.1). Alors,

1. Le pNB -QMLE est fortement consistant $\widehat{\theta}_r \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$ pour tout $r > 0$.

2. $\sqrt{n} \left(\widehat{\theta}_r - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, J_r^{-1} I_r J_r^{-1})$ pour tout $r > 0$, où les matrices I_r et J_r sont telles que

$$I_r = E \left(\frac{\alpha_0 (1 - \alpha_0) X_{t-1} + \sigma_0^2}{\lambda_t^2(\theta_0) (r + \lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right) \text{ et}$$

$$J_r = E \left(\frac{1}{\lambda_t(\widehat{\theta}_r) (r + \lambda_t(\widehat{\theta}_r))} \frac{\partial \lambda_t(\widehat{\theta}_r)}{\partial \theta} \frac{\partial \lambda_t(\widehat{\theta}_r)}{\partial \theta'} \right).$$

et la matrice de variance asymptotique estimée est telle que

$$\begin{aligned} V(\widehat{\theta}_r) &\simeq n^{-1} \widehat{J}_r^{-1} \widehat{I}_r \widehat{J}_r^{-1} \\ &= \left(\sum_{t=1}^n \frac{1}{\widetilde{\lambda}_t(\widehat{\theta}_r) (r + \widetilde{\lambda}_t(\widehat{\theta}_r))} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_r)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_r)}{\partial \theta'} \right)^{-1} \\ &\quad \times \left(\sum_{t=1}^n \frac{\widehat{\alpha}_r (1 - \widehat{\alpha}_r) X_{t-1} + \widehat{\sigma}_r^2}{\widetilde{\lambda}_t^2(\widehat{\theta}_r) (r + \widetilde{\lambda}_t(\widehat{\theta}_r))^2} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_r)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_r)}{\partial \theta'} \right) \\ &\quad \times \left(\sum_{t=1}^n \frac{1}{\widetilde{\lambda}_t(\widehat{\theta}_r) (r + \widetilde{\lambda}_t(\widehat{\theta}_r))} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_r)}{\partial \theta} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_r)}{\partial \theta'} \right)^{-1}. \end{aligned}$$

INAR avec innovation poissonnienne

Lorsque $\varepsilon_t \sim \mathcal{P}(w_0)$, ce qui veut dire $\sigma^2 = w_0$. Donc,

$$I_r = E \left(\frac{\alpha_0 (1 - \alpha_0) X_{t-1} + w_0}{\lambda_t^2(\theta_0) (r + \lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right)$$

Ainsi, la matrice de variance estimée est comme suit

$$\begin{aligned}
 V(\hat{\theta}_r) &\simeq \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_r) (r + \tilde{\lambda}_t(\hat{\theta}_r))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right)^{-1} \\
 &\times \left(\frac{\hat{\alpha}_r (1 - \hat{\alpha}_r) X_{t-1} + \hat{w}_0}{\tilde{\lambda}_t^2(\hat{\theta}_r) (r + \tilde{\lambda}_t(\hat{\theta}_r))^2} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right) \\
 &\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_r) (r + \tilde{\lambda}_t(\hat{\theta}_r))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right)^{-1}.
 \end{aligned}$$

INAR avec innovation géométrique

Lorsque $\varepsilon_t \sim \mathcal{G}\left(\frac{1}{w_0}\right)$, ce qui veut dire $\sigma^2 = w_0(w_0 - 1)$. Donc,

$$I_r = E \left(\frac{\alpha_0 (1 - \alpha_0) X_{t-1} + w_0 (w_0 - 1)}{\lambda_t^2(\theta_0) (r + \lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right)$$

Ainsi, la matrice de variance asymptotique estimée est donnée par

$$\begin{aligned}
 V(\hat{\theta}_r) &\simeq \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_r) (r + \tilde{\lambda}_t(\hat{\theta}_r))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right)^{-1} \\
 &\times \left(\sum_{t=1}^n \frac{\hat{\alpha}_r (1 - \hat{\alpha}_r) X_{t-1} + \hat{w}_r (\hat{w}_r - 1)}{\tilde{\lambda}_t^2(\hat{\theta}_r) (r + \tilde{\lambda}_t(\hat{\theta}_r))^2} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right) \\
 &\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_r) (r + \tilde{\lambda}_t(\hat{\theta}_r))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right)^{-1}.
 \end{aligned}$$

INAR avec innovation binomiale négative₂

Lorsque $\varepsilon_t \sim \mathcal{NB}\left(r, \frac{r}{r+w_0}\right)$, ce qui veut dire $\sigma^2 = \frac{1}{r}w_0(r+w_0)$. Donc,

$$I_r = E \left(\frac{\alpha_0(1-\alpha_0)X_{t-1} + \frac{1}{r}w_0(r+w_0)}{\lambda_t^2(\theta_0)(r+\lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right).$$

Alors, la matrice de variance théorique estimée est telle que

$$\begin{aligned} V(\hat{\theta}_r, r) &\simeq \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_r)(r+\tilde{\lambda}_t(\hat{\theta}_r))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right)^{-1} \\ &\times \left(\sum_{t=1}^n \frac{\hat{\alpha}_r(1-\hat{\alpha}_r)X_{t-1} + \frac{1}{r}\hat{w}_r(r+\hat{w}_r)}{\tilde{\lambda}_t^2(\hat{\theta}_r)(r+\tilde{\lambda}_t(\hat{\theta}_r))^2} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right) \\ &\times \left(\sum_{t=1}^n \frac{1}{\tilde{\lambda}_t(\hat{\theta}_r)(r+\tilde{\lambda}_t(\hat{\theta}_r))} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta} \frac{\partial \tilde{\lambda}_t(\hat{\theta}_r)}{\partial \theta'} \right)^{-1}. \end{aligned}$$

4.5 Estimation en quatre étapes : estimateur du QML binomial négatif en deux étapes (2SNB-QML)

Considérons le $2^{\text{ème}}$ cas du modèle (1.6.1)-(1.5.2) pour lequel nous étudions la procédure en quatre étapes susmentionnée *i) -iv)*. Ici, les deuxième et quatrième étapes sont décrites en plus de détails. En vertu de l'hypothèse GLM (1.5.1)-(1.5.2), si nous posons

$$u_t = (X_t - \lambda_t)^2 - E((X_t - \lambda_t)^2 | \mathcal{F}_{t-1}) = (X_t - \lambda_t)^2 - \left(1 + \frac{1}{r_0} \lambda_t\right) \lambda_t,$$

alors $E(u_t | \mathcal{F}_{t-1}) = 0$ et

$$\frac{(X_t - \lambda_t(\theta_0))^2 - \lambda_t(\theta_0)}{\lambda_t^2(\theta_0)} = \gamma_0 + \frac{u_t}{\lambda_t^2(\theta_0)}, \quad (4.5.1)$$

où $\gamma_0 = \frac{1}{r_0}$. Or la régression (4.5.1) n'est pas encore prête pour estimer γ_0 (et alors r_0) car le

régressant $\frac{(X_t - \lambda_t(\theta_0))^2 - \lambda_t(\theta_0)}{\lambda_t^2(\theta_0)}$ dépend du paramètre θ_0 qui est inconnu et donc elle est

non observable. Si une estimation consistante de θ_0 , disons $\hat{\theta}$, est disponible, nous pouvons former la régression modifiée (régression observable) suivante

$$\frac{(X_t - \hat{\lambda}_t)^2 - \hat{\lambda}_t}{\hat{\lambda}_t^2} = \gamma_0 + \frac{u_t}{\hat{\lambda}_t^2}, \quad (4.5.2)$$

à partir de laquelle une estimation consistante de r_0 , \hat{r} , l'inverse de l'estimation des moindres carrés pondérés $\hat{\gamma}$ de γ_0 , est donnée par

$$\hat{r} = \left(\frac{1}{n} \sum_{t=1}^n \frac{\left((X_t - \hat{\lambda}_t)^2 - \hat{\lambda}_t \right)}{\hat{\lambda}_t^2} \right)^{-1}, \quad \hat{\gamma} = \hat{r}^{-1}, \quad (4.5.3)$$

où $\hat{\lambda}_t = \tilde{\lambda}_t(\hat{\theta})$. Notons que l'estimation \hat{r} que nous utilisons ici est une adaptation dynamique de l'estimation proposée par Gourieroux et al (1984b) dans le contexte de la régression

binomiale négative statique. Maintenant, avec (4.5.3) l'algorithme suivant résume l'approche en quatre étapes *i) -iv)* décrite ci-dessus.

Algorithme 4.5.1 (NB-QMLE en deux étapes)

Etant donné un $r^* > 0$ fixé, l'estimateur du QML binomial négatif en deux étapes (2SNB-QMLE) de (θ_0, r_0) dans (1.6.1)-(1.5.2) se compose d'un quadruplet $(\hat{\theta}_{r^*}, \hat{r}_1, \hat{\theta}_{\hat{r}_1}, \hat{r}_2)$, qui est décrit par les étapes suivantes :

Étape 1 Posons $\hat{\theta}_{r^*} = \arg \max_{\theta \in \Theta} \tilde{L}_{n, r^*}(\theta)$, une solution du problème (4.2.1) tout en remplaçant r par r^* . Soit $\hat{\lambda}_{1t} = \tilde{\lambda}_t(\hat{\theta}_{r^*})$, $(1 \leq t \leq n)$.

Étape 2 Posons $\hat{\gamma}_1 = \frac{1}{n} \sum_{t=1}^n \frac{(X_t - \hat{\lambda}_{1t})^2 - \hat{\lambda}_{1t}}{\hat{\lambda}_{1t}^2}$, donc $\hat{r}_1 = \hat{\gamma}_1^{-1}$.

Étape 3 Soit $\hat{\theta}_{\hat{r}_1} = \arg \max_{\theta \in \Theta} \tilde{L}_{n, \hat{r}_1}(\theta)$ une solution du problème (4.2.1) tout en remplaçant le générique r par \hat{r}_1 . nous obtenons $\hat{\lambda}_{2t} = \tilde{\lambda}_t(\hat{\theta}_{\hat{r}_1})$, $(1 \leq t \leq n)$.

Étape 4 Soit $\hat{\gamma}_2 = \frac{1}{n} \sum_{t=1}^n \frac{(X_t - \hat{\lambda}_{2t})^2 - \hat{\lambda}_{2t}}{\hat{\lambda}_{2t}^2}$, ainsi il résulte que $\hat{r}_2 = \hat{\gamma}_2^{-1}$.

4.6 Propriétés asymptotiques du 2SNB-QMLE

Pour obtenir les propriétés asymptotiques du quadruplet $(\hat{\theta}_{r^*}, \hat{r}_1, \hat{\theta}_{\hat{r}_1}, \hat{r}_2)$, notons d'abord que $\hat{\theta}_{r^*}$ n'est autre que le profil NB-QMLE proposé dans la section 4.2 dont les propriétés asymptotiques ont été données par le théorème 4.3.1 et le théorème 4.3.2. Il reste donc à étudier le triplet $(\hat{r}_1, \hat{\theta}_{\hat{r}_1}, \hat{r}_2)$, dont les propriétés asymptotiques sont données par le résultat suivant.

Théorème 4.6.1 (Aknouche, Bendjeddou et Touche, 2018) *Sous (1.6.1), (1.5.1)-(1.5.2) et **A0-A10**, nous avons*

$$\widehat{r}_1 \xrightarrow[n \rightarrow \infty]{p.s.} r_0, \quad (4.6.1a)$$

$$\sqrt{n}(\widehat{\gamma}_1 - \gamma_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, E \left(\frac{\left((X_t - \lambda_t(\theta_0))^2 - \left(\lambda_t(\theta_0) + \frac{1}{r_0} \lambda_t^2(\theta_0) \right) \right)^2}{\lambda_t^4(\theta_0)} \right) \right), \widehat{\gamma}_2 \stackrel{A.D}{=} \widehat{\gamma}_1, \quad (4.6.1b)$$

$$\widehat{\theta}_{\widehat{r}_1} \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0, \quad (4.6.1c)$$

$$\sqrt{n}(\widehat{\theta}_{\widehat{r}_1} - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N \left(0, \frac{1}{r_0} J_{r_0}^{-1} \right), \quad (4.6.1d)$$

Où $\stackrel{A.D}{=}$ représente l'égalité en distribution asymptotique.

Remarque 4.6.1 Quelques conclusions générales peuvent être tirées :

1°\ . La consistance forte de $\widehat{\theta}_{\widehat{r}_1}$ découle directement de la consistance forte de $\widehat{\theta}_r$ (pour tout $r > 0$) et \widehat{r}_1 .

2°\ . L'estimation de la troisième étape $\widehat{\theta}_{\widehat{r}_1}$ est clairement asymptotiquement plus efficace que l'estimation de la première étape $\widehat{\theta}_{r^*}$.

3°\ . Aucune hypothèse de moment supplémentaire, à part celles requises par **A0-A10**, n'est nécessaire pour la consistance et la normalité asymptotique de $\widehat{\gamma}_1$. D'autres méthodes d'estimation de γ sont disponibles (e.g. Christou et Fokianos, 2014), mais elles peuvent impliquer des conditions d'ordre supérieur.

4°\ . La distribution asymptotique de \widehat{r}_1 est une *distribution normale réciproque*, qui est bimodale et n'admet pas un moment d'ordre un.

5°\ . Comme $\widehat{\gamma}_1$ et $\widehat{\gamma}_2$ ont la même distribution asymptotique, l'**Étape 4** est facultative et peut être omise. Cependant, à l'égard d'échantillons finis, nous la gardons ici parce qu'elle permet de réestimer r_0 en utilisant $\widehat{\lambda}_{2t}$ et donc $\widehat{\theta}_{\widehat{r}_1}$, qui est asymptotiquement plus efficace que $\widehat{\theta}_r$ que nous avons utilisé à l'**Étape 2**.

6°. Une estimation consistante de la variance asymptotique $\frac{1}{r_0} J_{r_0}^{-1}$, de la troisième étape d'estimation $\widehat{\theta}_{\widehat{r}_1}$, est

$$\frac{1}{\widehat{r}_2} \widehat{J}_{\widehat{r}_2}^{-1}, \quad (4.6.2)$$

où \widehat{J}_r est donnée par (4.3.6). Et vu que $I_r = J_r$, alors (4.3.7) peut également être utilisée.

7°. Une estimation consistante de la variance asymptotique de $\widehat{\gamma}_2$ dans (4.6.1b) est

$$\frac{1}{n} \sum_{t=1}^n \frac{\left(\left(X_t - \lambda_t(\widehat{\theta}_{\widehat{r}_1}) \right)^2 - \left(\lambda_t(\widehat{\theta}_{\widehat{r}_1}) + \frac{1}{r_0} \lambda_t^2(\widehat{\theta}_{\widehat{r}_1}) \right) \right)^2}{\lambda_t^4(\widehat{\theta}_{\widehat{r}_1})}. \quad (4.6.3)$$

8°. Les sorties de la méthode *2SNB-QMLE* sont $\widehat{r}_2 = (\widehat{\gamma}_2)^{-1}$ et $\widehat{\theta}_{\widehat{r}_1}$.

4.7 Application au modèle *INGARCH* binomial négatif₂

Proposition 4.7.1 (Aknouche, Bendjeddou et Touche, 2018) Soit $\theta_0 \in \Theta$ un sous ensemble compact de Θ_2 et X une solution stationnaire de (1.3.5). Alors,

1. Le *pNB-QMLE* est fortement consistant $\widehat{\theta}_r \xrightarrow[n \rightarrow \infty]{p.s.} \theta_0$ pour tout $r > 0$.

2. $\sqrt{n} \left(\widehat{\theta}_r - \theta_0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, J_r^{-1} I_r J_r^{-1})$ pour tout $r > 0$, où les matrices I_r et J_r sont

telles que

$$I_r = \frac{1}{r_0} E \left(\frac{1}{\lambda_t(\theta_0) (r + \lambda_t(\theta_0))} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right) = \frac{1}{r_0} J_r.$$

Dans ce cas, la matrice de variance asymptotique est telle que

$$V \left(\widehat{\theta}_{r_2}, r_2 \right) \simeq n^{-1} \frac{1}{\widehat{r}_2} \widehat{J}_{\widehat{r}_2}^{-1} = \frac{1}{\widehat{r}_2} \left(\sum_{t=1}^n \frac{1}{\widetilde{\lambda}_t(\widehat{\theta}_{\widehat{r}_2}) \left(\widehat{r}_2 + \widetilde{\lambda}_t(\widehat{\theta}_{\widehat{r}_2}) \right)} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_{\widehat{r}_2})}{\partial \theta} \frac{\partial \widetilde{\lambda}_t(\widehat{\theta}_{\widehat{r}_2})}{\partial \theta'} \right)^{-1}.$$

4.8 Preuves des principaux résultats

Preuve du Théorème 4.3.1

En suivant l'approche de Wald, la preuve du Théorème 4.3.1 repose sur les trois lemmes suivants.

Lemme 4.8.1 Sous **A1-A2**

$$\lim_{n \rightarrow \infty} \sup_{\theta \in \Theta} \left| L_{n,r}(\theta) - \tilde{L}_{n,r}(\theta) \right| = 0, \quad p.s.$$

Preuve En utilisant l'inégalité $\log(x) \leq x - 1$, le fait que $\tilde{\lambda}_t(\theta) > 0$, les hypothèses **A1-A2** et puis le lemme de Césaro il vient que

$$\begin{aligned} \sup_{\theta \in \Theta} \left| L_{n,r}(\theta) - \tilde{L}_{n,r}(\theta) \right| &= \frac{1}{n} \sup_{\theta \in \Theta} \left| \sum_{t=1}^n \left(\log \left(\frac{r + \lambda_t(\theta)}{r + \tilde{\lambda}_t(\theta)} \right) + X_t \log \left(\frac{\tilde{\lambda}_t(\theta)(r + \lambda_t(\theta))}{\lambda_t(\theta)(r + \tilde{\lambda}_t(\theta))} \right) \right) \right| \\ &= \frac{1}{n} \sup_{\theta \in \Theta} \left| \sum_{t=1}^n \left(\log \left(\frac{\lambda_t(\theta) - \tilde{\lambda}_t(\theta)}{r + \tilde{\lambda}_t(\theta)} + 1 \right) + X_t \log \left(r \frac{\tilde{\lambda}_t(\theta) - \lambda_t(\theta)}{\lambda_t(\theta)(r + \tilde{\lambda}_t(\theta))} + 1 \right) \right) \right| \\ &\leq \frac{1}{n} \sum_{t=1}^n \left(\frac{1}{r} \sup_{\theta \in \Theta} \left| \lambda_t(\theta) - \tilde{\lambda}_t(\theta) \right| + X_t \sup_{\theta \in \Theta} \left| \tilde{\lambda}_t(\theta) - \lambda_t(\theta) \right| \frac{r}{cr} \right) \\ &= \frac{1}{n} \sum_{t=1}^n \left(\frac{1}{r} a_t + \frac{1}{c} X_t a_t \right) \xrightarrow[n \rightarrow \infty]{p.s.} 0. \end{aligned}$$

■

Lemme 4.8.2 Sous **A0-A4**,

- i) $E(l_{1,r}(\theta_0)) < \infty$.
- ii) $E(l_{1,r}(\theta_0)) \geq E(l_{1,r}(\theta))$ pour tout $\theta \in \Theta$.
- iii) $E(l_{1,r}(\theta)) = E(l_{1,r}(\theta_0)) \Rightarrow \theta = \theta_0$.

Preuve Sous **A1**, les variables aléatoires $\log \left(\frac{r}{r + \lambda_t(\theta)} \right)$ et $\log \left(\frac{\lambda_t(\theta)}{r + \lambda_t(\theta)} \right)$ sont bornées.

Par conséquent, elles admettent des moments finis de tout ordre. A l'aide de **A3** et les

inégalités de Jensen et Hölder, il s'ensuit que

$$\begin{aligned}
 |E(l_{1,r}(\theta_0))| &\leq E(|l_{1,r}(\theta_0)|) \leq E\left(\left|\log\left(\frac{r}{r+\lambda_t(\theta_0)}\right)\right|\right) + E\left(\left|X_t \log\left(\frac{\lambda_t(\theta_0)}{r+\lambda_t(\theta_0)}\right)\right|\right) \\
 &\leq E\left(\left|\log\left(\frac{r}{r+\lambda_t(\theta_0)}\right)\right|\right) + (E(X_t^\delta))^{1/\delta} \left(E\left(\left|\log\frac{\lambda_t(\theta_0)}{r+\lambda_t(\theta_0)}\right|^{\frac{\delta}{\delta-1}}\right)\right)^{\frac{\delta-1}{\delta}} < \infty.
 \end{aligned} \tag{4.8.1}$$

D'autre part, en utilisant à nouveau l'inégalité $\log(x) \leq x - 1$, nous avons

$$\begin{aligned}
 E(l_{1,r}(\theta) - l_{1,r}(\theta_0)) &= E\left(r \log\left(\frac{r+\lambda_t(\theta_0)}{r+\lambda_t(\theta)}\right) + X_t \log\left(\frac{\lambda_t(\theta)(r+\lambda_t(\theta_0))}{\lambda_t(\theta_0)(r+\lambda_t(\theta))}\right)\right) \\
 &\leq rE\left(\left(\frac{r+\lambda_t(\theta_0)}{r+\lambda_t(\theta)} - 1\right) + X_t \left(\frac{\lambda_t(\theta)(r+\lambda_t(\theta_0))}{\lambda_t(\theta_0)(r+\lambda_t(\theta))} - 1\right)\right) \\
 &= rE\left(\frac{\lambda_t(\theta_0) - \lambda_t(\theta)}{r+\lambda_t(\theta)}\right) + E\left(r \frac{X_t}{\lambda_t(\theta_0)} \frac{\lambda_t(\theta) - \lambda_t(\theta_0)}{r+\lambda_t(\theta)}\right) \\
 &= rE\left(\frac{\lambda_t(\theta_0) - \lambda_t(\theta)}{r+\lambda_t(\theta)} + \frac{\lambda_t(\theta) - \lambda_t(\theta_0)}{r+\lambda_t(\theta)}\right) = 0,
 \end{aligned} \tag{4.8.2}$$

Par (4.8.1) et (4.8.2) il s'ensuit que $E(l_{1,r}(\theta) - l_{1,r}(\theta_0)) \in [-\infty, 0]$ alors $E(l_{1,r}(\theta)) < E(l_{1,r}(\theta_0))$ pour tout $\theta \neq \theta_0$. Enfin, l'inégalité (4.8.2) se réduit à une égalité si et seulement si

$$rE\left(\log\left(\frac{r+\lambda_t(\theta_0)}{r+\lambda_t(\theta)}\right) + X_t \log\left(\frac{\lambda_t(\theta)(r+\lambda_t(\theta_0))}{\lambda_t(\theta_0)(r+\lambda_t(\theta))}\right)\right) = 0,$$

qui tient si et seulement si $\lambda_t(\theta) = \lambda_t(\theta_0)$ et ensuite, par l'hypothèse d'identifiabilité **A4**, si et seulement si $\theta = \theta_0$. ■

Lemme 4.8.3 Sous **A0-A5**, Il existe pour tout $\theta \neq \theta_0$ un voisinage $V(\theta)$ tel que

$$\limsup_{n \rightarrow \infty} \sup_{\theta^* \in V(\theta)} \tilde{L}_{n,r}(\theta^*) < \limsup_{n \rightarrow \infty} \tilde{L}_{n,r}(\theta_0) \quad p.s. \tag{4.8.3}$$

Preuve Pour tout $\bar{\theta} \in \Theta$ et $k \in \mathbb{N}^*$, soit $V_k(\bar{\theta})$ la boule ouverte de centre $\bar{\theta}$ et de rayon

$1/k$. Puisque $\sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} l_{t,r}(\theta)$ est une fonction mesurable de la suite $\{X_t, t \in \mathbb{Z}\}$, qui est strictement stationnaire et ergodique sous **A0**, alors $\left\{ \sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} l_{t,r}(\theta), t \in \mathbb{Z} \right\}$ est également strictement stationnaire et ergodique où par le Lemme 4.8.2 $E \left(\sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} l_{t,r}(\theta) \right) \in]-\infty, +\infty[$. Par conséquent, compte tenu du Lemme 4.8.1 et du théorème ergodique (Billingsley, 2008), il s'ensuit que

$$\limsup_{n \rightarrow \infty} \sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} \tilde{L}_{n,r}(\theta) = \limsup_{n \rightarrow \infty} \sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} L_{n,r}(\theta) \leq E \left(\sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} l_{1,r}(\theta) \right).$$

Par le théorème de Beppo-Levi $E \left(\sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} l_{1,r}(\theta) \right)$ converge vers $E(l_{1,r}(\bar{\theta}))$ lorsque $k \rightarrow \infty$. Par conséquent, (4.8.3) résulte du Lemme 4.8.2, ii). ■

En vue des Lemmes 4.8.1-4.8.3, nous avons montré qu'il existe pour tout $\bar{\theta} \neq \theta_0$ un voisinage $V(\bar{\theta})$ tel que

$$\limsup_{n \rightarrow \infty} \sup_{\theta \in V_k(\bar{\theta}) \cap \Theta} \tilde{L}_{n,r}(\theta) < \limsup_{n \rightarrow \infty} \tilde{L}_{n,r}(\theta_0) = \limsup_{n \rightarrow \infty} L_{n,r}(\theta_0) = E(l_{1,r}(\theta_0)).$$

Ainsi, à partir d'arguments standards, la démonstration du théorème 4.3.1 est achevée en utilisant l'hypothèse **A5** de compacité de Θ .

Preuve du Théorème 4.3.2

De **A7** et du théorème 4.3.1, nous savons que $\hat{\theta}_r$ ne peut pas être à la frontière de Θ pour n

suffisamment grand. Par conséquent, un développement de Taylor de $\frac{\partial L_{n,r}(\hat{\theta}_r)}{\partial \theta}$ à θ_0 donne

$$\begin{aligned} 0 &= \sqrt{n} \frac{\partial \tilde{L}_{n,r}(\hat{\theta}_r)}{\partial \theta} \\ &= \sqrt{n} \frac{\partial L_{n,r}(\hat{\theta}_r)}{\partial \theta} + \sqrt{n} \left(\frac{\partial \tilde{L}_{n,r}(\theta)}{\partial \theta} - \frac{\partial L_{n,r}(\theta)}{\partial \theta} \right) \\ &= \sqrt{n} \frac{\partial L_{n,r}(\theta_0)}{\partial \theta} + \sqrt{n} \frac{\partial^2 L_{n,r}(\theta^*)}{\partial \theta \partial \theta'} (\hat{\theta}_r - \theta_0) + \sqrt{n} \left(\frac{\partial \tilde{L}_{n,r}(\theta)}{\partial \theta} - \frac{\partial L_{n,r}(\theta)}{\partial \theta} \right) \end{aligned} \quad (4.8.4)$$

pour un certain θ^* entre $\hat{\theta}_r$ et θ_0 . En vue de (4.8.3), la démonstration du théorème 4.3.2 est basée sur les trois Lemmes suivants. Le Lemme 4.8.4 montre que le dernier terme dans (4.8.4) est p.s. négligeable lorsque $n \rightarrow \infty$. Le Lemme 4.8.5 établit la convergence en loi du premier terme de (4.8.4) en utilisant le théorème central limite des martingales tandis que le Lemme 4.8.6 montre la convergence vers la matrice du le second terme de (4.8.4).

Lemme 4.8.4 *Sous A0-A10*

$$\sqrt{n} \sup_{\theta \in \Theta} \left\| \frac{\partial \tilde{L}_{n,r}(\theta)}{\partial \theta} - \frac{\partial L_{n,r}(\theta)}{\partial \theta} \right\| \xrightarrow[n \rightarrow \infty]{p.s.} 0.$$

Preuve En utilisant **A2** et **A6**, il s'ensuit que

$$\begin{aligned} \sqrt{n} \sup_{\theta \in \Theta} \left\| \frac{\partial \tilde{L}_{n,r}(\theta)}{\partial \theta} - \frac{\partial L_{n,r}(\theta)}{\partial \theta} \right\| &= \frac{1}{\sqrt{n}} \sup_{\theta \in \Theta} \left\| \sum_{t=1}^n \left[\frac{\partial}{\partial \theta} \left(\log \left(\frac{r}{r + \tilde{\lambda}_t(\theta)} \right) + X_t \log \left(\frac{\tilde{\lambda}_t(\theta)}{r + \tilde{\lambda}_t(\theta)} \right) \right) \right. \right. \\ &\quad \left. \left. - \frac{\partial}{\partial \theta} \left(\log \left(\frac{r}{r + \lambda_t(\theta)} \right) + X_t \log \left(\frac{\lambda_t(\theta)}{r + \lambda_t(\theta)} \right) \right) \right] \right\| \\ &\leq \frac{1}{\sqrt{n}} \sum_{t=1}^n \left(c_t + a_t d_t + X_t \left(\frac{c_t}{cr} + \frac{(a_t + b_t) d_t}{c^2 r^2} \right) \right) \xrightarrow[n \rightarrow \infty]{p.s.} 0. \end{aligned}$$

Lemme 4.8.5 *Sous A8-A9*,

$$\sqrt{n} \frac{\partial L_{n,r}(\theta_0)}{\partial \theta} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0, I_r).$$

Preuve Il est clair que $\left\{ \sqrt{n} \frac{\partial L_{n,r}(\theta_0)}{\partial \theta}, t \in \mathbb{Z} \right\}$ est une martingale par rapport à $\{\mathcal{F}_t, t \in \mathbb{Z}\}$,

où

$$\begin{aligned} \sqrt{n} \frac{\partial L_{n,r}(\theta_0)}{\partial \theta} &= \sum_{t=1}^n \frac{1}{\sqrt{n}} \frac{\partial l_{t,r}(\theta_0)}{\partial \theta}, \\ \frac{\partial l_{t,r}(\theta_0)}{\partial \theta} &= \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{X_t - \lambda_t(\theta_0)}{\lambda_t(\theta_0) (1 + \lambda_t(\theta_0))}. \end{aligned}$$

Du **A8-A9**, nous avons

$$E \left(\frac{\partial l_{t,r}(\theta_0)}{\partial \theta} \frac{\partial l_{t,r}(\theta_0)}{\partial \theta'} \right) = E \left(\frac{v_t(\theta_0)}{\lambda_t^2(\theta_0) (1 + \lambda_t(\theta_0))^2} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta'} \right) = I_r.$$

Ainsi, le lemme 4.8.5 découle du théorème central limite des martingales (voir Billingsley, 2008). ■

Lemme 4.8.6 *Sous A8-A10,*

$$\frac{\partial^2 L_{n,r}(\theta^*)}{\partial\theta\partial\theta'} \xrightarrow[n \rightarrow \infty]{p.s.} J_r.$$

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 $V(\theta_0)$ défini par **A10**. Supposons que n est assez grand pour que θ^* appartienne à $V_k(\theta_0)$. Donc par stationnarité et ergodicité de

$$\left\{ \sup_{\theta \in V_k(\theta_0)} \left| \frac{\partial^2 l_{t,r}(\theta)}{\partial\theta_i \partial\theta_j} - E \left(\frac{\partial^2 l_{t,r}(\theta_0)}{\partial\theta_i \partial\theta_j} \right) \right| \right\},$$

nous avons

$$\begin{aligned} \left| \frac{\partial^2 L_{n,r}(\theta^*)}{\partial\theta_i \partial\theta_j} - J_r(i,j) \right| &= \left| \frac{\partial^2 L_{n,r}(\theta^*)}{\partial\theta_i \partial\theta_j} - E \left(\frac{\partial^2 L_{n,r}(\theta_0)}{\partial\theta_i \partial\theta_j} \right) \right| \\ &= \frac{1}{n} \left| \sum_{t=1}^n \frac{\partial^2 l_{t,r}(\theta^*)}{\partial\theta_i \partial\theta_j} - E \left(\frac{\partial^2 l_{t,r}(\theta_0)}{\partial\theta_i \partial\theta_j} \right) \right| \\ &\leq \frac{1}{n} \sup_{\theta \in V_k(\theta_0)} \left| \sum_{t=1}^n \frac{\partial^2 l_{t,r}(\theta)}{\partial\theta_i \partial\theta_j} - E \left(\frac{\partial^2 l_{t,r}(\theta_0)}{\partial\theta_i \partial\theta_j} \right) \right| \\ &\leq \frac{1}{n} \sum_{t=1}^n \sup_{\theta \in V_k(\theta_0)} \left| \frac{\partial^2 l_{t,r}(\theta)}{\partial\theta_i \partial\theta_j} - E \left(\frac{\partial^2 l_{t,r}(\theta_0)}{\partial\theta_i \partial\theta_j} \right) \right| \\ &\xrightarrow[n \rightarrow \infty]{p.s.} E \left(\sup_{\theta \in V_k(\theta_0)} \left| \frac{\partial^2 l_{t,r}(\theta)}{\partial\theta_i \partial\theta_j} - E \left(\frac{\partial^2 l_{t,r}(\theta_0)}{\partial\theta_i \partial\theta_j} \right) \right| \right). \end{aligned}$$

Au vu de **A10**, le théorème de convergence dominée de Lebesgue entraîne que

$$\begin{aligned} \lim_{k \rightarrow \infty} E \left(\sup_{\theta \in V_k(\theta_0)} \left| \frac{\partial^2 l_{t,r}(\theta)}{\partial\theta_i \partial\theta_j} - E \left(\frac{\partial^2 l_{t,r}(\theta_0)}{\partial\theta_i \partial\theta_j} \right) \right| \right) &= E \left(\lim_{k \rightarrow \infty} \sup_{\theta \in V_k(\theta_0)} \left| \frac{\partial^2 l_{t,r}(\theta)}{\partial\theta_i \partial\theta_j} - E \left(\frac{\partial^2 l_{t,r}(\theta_0)}{\partial\theta_i \partial\theta_j} \right) \right| \right) \\ &= 0, \end{aligned}$$

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

Preuve du Théorème 4.6.1

i) Preuve de (4.6.1a) Il suffit de prouver la consistance forte de $\hat{\gamma}$. Depuis (4.5.2) et (4.5.3) nous avons

$$\begin{aligned}\hat{\gamma} - \gamma_0 &= \frac{1}{n} \sum_{t=1}^n \frac{u_t}{\hat{\lambda}_t^2} \\ &= \frac{1}{n} \sum_{t=1}^n \frac{u_t}{\lambda_t^2} + \frac{1}{n} \sum_{t=1}^n \frac{u_t}{\hat{\lambda}_t^2} \left(\frac{1}{\hat{\lambda}_t^2} - \frac{1}{\lambda_t^2} \right).\end{aligned}\quad (4.8.5)$$

Par le théorème ergodique, le premier terme du côté droit de (4.8.5) satisfait le résultat limitatif suivant

$$\frac{1}{n} \sum_{t=1}^n \frac{u_t}{\lambda_t^2} \xrightarrow[n \rightarrow \infty]{p.s.} E \left(\frac{u_t}{\lambda_t^2} \right) = E \left(\frac{1}{\lambda_t^2} E(u_t | \mathcal{F}_{t-1}) \right) = 0.$$

Il reste donc à montrer que

$$\frac{1}{n} \sum_{t=1}^n \frac{u_t}{\hat{\lambda}_t^2} \left(\frac{1}{\hat{\lambda}_t^2} - \frac{1}{\lambda_t^2} \right) = o_{p.s.}(1).\quad (4.8.6)$$

En utilisant une expansion de Taylor de $\frac{1}{\lambda_t^2(\hat{\theta}_r)}$ autour de θ_0 , nous avons

$$\begin{aligned}\frac{1}{\hat{\lambda}_t^2} - \frac{1}{\lambda_t^2} &= \frac{1}{\lambda_t^2(\hat{\theta}_r)} - \frac{1}{\lambda_t^2(\theta_0)} \\ &= -\frac{2}{\lambda_t^3(\theta^*)} \frac{\partial \lambda_t(\theta^*)}{\partial \theta'} (\hat{\theta}_r - \theta_0),\end{aligned}$$

Où θ^* est comprise entre $\hat{\theta}_r$ et θ_0 . Ainsi, (4.8.6) résulte de **A1**, **A10**, la consistance forte de $\hat{\theta}_r$ et le théorème ergodique.

ii) Preuve de (4.6.1b) Réécrivons (4.8.5) comme suit

$$\sqrt{n}(\hat{\gamma} - \gamma_0) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{u_t}{\lambda_t^2} + \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{u_t}{\hat{\lambda}_t^2} \left(\frac{1}{\hat{\lambda}_t^2} - \frac{1}{\lambda_t^2} \right).$$

Si nous montrons que

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{u_t}{\widehat{\lambda}_t^2} \left(\frac{1}{\widehat{\lambda}_t} - \frac{1}{\lambda_t^*} \right) = o_p(1), \quad (4.8.7)$$

alors (4.6.1b) découlerait du théorème central limite des martingales appliqué à la différence

$\{\mathcal{F}_t, t \in \mathbb{Z}\}$ -martingale $\left\{ \frac{u_t}{\lambda_t^2}, t \in \mathbb{Z} \right\}$. Maintenant, par une expansion de Taylor de $\frac{1}{\lambda_t^2(\widehat{\theta}_r)}$

autour de θ_0 , le côté gauche de (4.8.7) devient

$$\frac{-2 \left(\widehat{\theta}_r - \theta_0 \right)'}{\sqrt{n}} \sum_{t=1}^n \frac{u_t}{\widehat{\lambda}_t^2 \lambda_t^3(\theta^*)} \frac{\partial \lambda_t(\theta^*)}{\partial \theta},$$

et (4.8.7) découle des hypothèses **A1** et **A10**, la normalité asymptotique de $\sqrt{n} \left(\widehat{\theta}_r - \theta_0 \right)$, ce qui implique que

$$\widehat{\theta}_r - \theta_0 = n^{-1/2} O_p(1),$$

et le théorème ergodique.

iii) Preuve de (4.6.1c) Le résultat (4.6.1c) est une conséquence évidente de la forte consistance de $\widehat{\theta}_r$ (Voir (4.3.1)) pour tout $r > 0$.

iv) Preuve de (4.6.1d) De la consistance de \widehat{r}_1 et de la \sqrt{n} -consistance de $\widehat{\theta}_r$ pour tout $r > 0$, nous avons

$$\begin{aligned} \sqrt{n} \left(\widehat{\theta}_{\widehat{r}_1} - \theta_0 \right) &= \sqrt{n} \left(\widehat{\theta}_{r_0} - \theta_0 \right) + \sqrt{n} \left(\widehat{\theta}_{\widehat{r}_1} - \widehat{\theta}_{r_0} \right) \\ &= \sqrt{n} \left(\widehat{\theta}_{r_0} - \theta_0 \right) + o_p(1), \end{aligned}$$

alors, le résultat découle du théorème 4.3.2 tout en remplaçant r par r_0 (voir (4.3.3)) et en utilisant le fait que $I_{r_0} = J_{r_0}$. ■

Chapitre 5

Comparaison des méthodes du QML dans le cadre du modèle $INGAR$

5.1 Introduction

Dans ce chapitre, nous allons comparer les estimateurs $NB-QML$ et $P-QMLE$ au vu de leur efficacité asymptotique relative ensuite pour des échantillons finis à l'aide de quelques expériences de simulation et ce pour certains cas spécifiques bien connus de (1.6.1) ; à savoir le modèle $INGARCH$ poissonien, le modèle $INGARCH$ binomial négatif de type 1, le modèle $INGARCH$ binomial négatif de type 2 ainsi que le modèle $INAR(1)$. Nous comparons également ces estimateurs sur certaines fonctions de lien GLM particulières de (1.5.1). Autres cas particuliers du modèle $INGAR$ sont comparés aussi tels que les modèles $INGARCH$ avec distributions conditionnelles double Poisson et Poisson généralisée dont les détails sont disponibles dans le support online de Aknouche et *al* (2018b).

5.2 Comparaison sous des modèles *INGARCH* (1, 1)

5.2.1 Comparaison vis-à-vis de l'efficacité asymptotique relative

En considérant le cas du modèle *INGARCH* poissonien (1.3.1), pour lequel la fonction de lien est la fonction "identité" donnée par

$$\text{Var}(X_t | \mathcal{F}_{t-1}) = E(X_t | \mathcal{F}_{t-1}). \quad (5.2.1)$$

Ainsi, le *P-QMLE* de (1.3.1) se réduit à l'estimateur du maximum de vraisemblance, qui est asymptotiquement efficace et est alors plus asymptotiquement efficace que le *pNB-QMLE* quant à l'estimateur *2SNB-QML*, donné par la section (4.5), il ne convient pas avec le présent cas vu que la deuxième étape de l'algorithme 4.5.1 est dérivée sous l'hypothèse GLM (1.5.1)-(1.5.2), qui est différente de la fonction de lien (5.2.1) caractérisant le modèle *INGARCH* poissonien (1.3.1).

Maintenant, lorsque le modèle *INGARCH* est *K*-binomial négatif, *NB_K-INGARCH*(*p*, *q*), la comparaison s'effectue par rapport aux différentes valeurs de *K* ; à savoir *K* = 1, *K* = 2 et *K* ∉ {1, 2}.

En observant la fonction de lien de ce modèle, (1.3.3), pour lequel

$$\begin{aligned} E(X_t | \mathcal{F}_{t-1}) &= \lambda_t, \\ \text{Var}(X_t | \mathcal{F}_{t-1}) &= E(X_t | \mathcal{F}_{t-1}) \left(1 + \frac{1}{r_0} (E(X_t | \mathcal{F}_{t-1}))^{K-1}\right), \end{aligned} \quad (5.2.2)$$

ce qui implique d'ailleurs que le processus est surdispersé vu que $\text{Var}(X_t | \mathcal{F}_{t-1}) > E(X_t | \mathcal{F}_{t-1})$, pour *K* = 1,

$$\text{Var}(X_t | \mathcal{F}_{t-1}) = \left(1 + \frac{1}{r_0}\right) E(X_t | \mathcal{F}_{t-1}), \quad (5.2.3)$$

on voit facilement que, sous la condition du Modèle Linéaire Généralisé poissonien 5.2.3, le *P-QMLE* est asymptotiquement efficace dans la classe de tous les estimateurs *QML* appartenant à la famille exponentielle linéaire. Ainsi, le *P-QMLE* est asymptotiquement plus efficace que le *pNB - QMLE* (voir Gourieroux et al, (1984a, 1984b) dans le cas des

modèles de régression statique à valeurs entières où l'adaptation au cas dynamique actuel est triviale). Notons enfin que, comme dans le cas *INGARCH* poissonien, le *2SNB-QMLE* donné par la section 4.5 est mal défini.

Au vu du modèle (1.3.3), cas où $K = 2$, la fonction de lien (5.2.2) se réduit clairement à la condition GLM (4.3.4) binomiale négative₂ (avec $\delta^2 = 1$), c'est-à-dire

$$\text{Var}(X_t | \mathcal{F}_{t-1}) = E(X_t | \mathcal{F}_{t-1}) \left(1 + \frac{1}{r_0} E(X_t | \mathcal{F}_{t-1})\right), \quad r_0 > 0, \quad (5.2.4)$$

sous laquelle le *2SNB-QMLE* est dérivé. Ainsi, aucun des estimateurs *pNB-QML* et *P-QML* n'est asymptotiquement supérieur à l'autre, à moins que r_0 ne soit connu. Dans ce cas, on peut prendre $r = r_0$ et le *pNB-QMLE*, $\hat{\theta}_{r_0}$, serait asymptotiquement efficace. Par exemple, considérons le modèle *INGARCH* géométrique qui est un cas particulier du modèle *NB₂-INGARCH*(1.3.5) dans lequel $r_0 = 1$, c'est-à-dire

$$X_t | \mathcal{F}_{t-1} \sim \mathcal{G}\left(\frac{1}{1+\lambda_t}\right).$$

Pour ce modèle, le *QMLE* (*G-QMLE*) géométrique, qui est un cas particulier du *pNB-QMLE* correspondant à $r = 1$, se réduit à l'estimateur du maximum de vraisemblance et est alors asymptotiquement efficace. Cependant, qu'on sache ou non, l'estimateur *2SNB-QML* a l'intéressante propriété d'être asymptotiquement efficace dans la classe de tous les estimateurs *QML* appartenant à la famille exponentielle linéaire (voir le théorème 4.6.1). Par conséquent, il est asymptotiquement plus efficace que le *P-QMLE*.

Enfin, il est intéressant de noter que lorsque $K \notin \{1, 2\}$, la fonction de lien (5.2.2) correspondante au modèle *NB_K-INGARCH* est différente de la condition du *GLM* poissonien (5.2.3) et de l'hypothèse binomiale négative₂ (5.2.4). Par conséquent, l'estimateur *2SNB-QML* est mal défini et aucun des *P-QMLE* et *pNB-QMLE* n'est asymptotiquement préférable que l'autre.

5.2.2 Comparaison en échantillons finis : étude de simulation

Nous examinons maintenant la performance en échantillons finis des estimateurs *NB-QML* proposés sur des séries simulées avec une taille d'échantillon $n = 1000$. Ces séries sont générées à partir de trois instances de (1.6.1), à savoir :

i) Le modèle *INGARCH*(1, 1) poissonien (1.3.1) avec le paramètre $\theta_0 = (2, 0.6, 0.3)'$ (cf. tableau 5.2.1).

ii) Le modèle *INGARCH*(1, 1) géométrique correspondant à (1.3.5) avec $r_0 = 1$ et $\theta_0 = (2, 0.6, 0.3)'$ (cf. tableau 5.2.2).

iii) Le modèle *INGARCH*(1, 1) binomial négatif₂ (1.3.5) avec les paramètres $r_0 = 3$ et $\theta_0 = (2, 0.6, 0.3)'$ (cf. tableau 5.2.3).

Trois estimateurs *QML* sont comparés sur ces modèles :

- i)** L'estimateur *QML* poissonien ($\hat{\theta}_P$, Ahmad et Francq, 2016) donné par (2.4.1),
- ii)** l'estimateur *QML* géométrique, $\hat{\theta}_1$, correspondant à (4.2.1) avec $r = 1$, et
- iii)** l'estimateur *QML* binomial négatif en profil, $\hat{\theta}_4$, donné par (4.2.1) avec $r = 4$. Pour le modèle *NB₂-INGARCH*(1, 1) (1.3.5) on exécute aussi l'approche d'estimation *NB-QMLE* en deux étapes ($\hat{\theta}_{r^*}, \hat{r}_1, \hat{\theta}_{\hat{r}_1}, \hat{r}_2$), décrite par l'algorithme 4.5.1.

Ces estimations sont calculées en utilisant 500 réplifications de Monte Carlo pour les trois modèles mentionnés.

Dans l'implémentation du *NB-QMLE* nous avons utilisé les mêmes dispositifs : la valeur du paramètre initial, $\theta^{(0)} = (\omega^{(0)}, \alpha^{(0)}, \beta^{(0)})'$, de la routine d'optimisation non linéaire (4.2.1) est fixée à la valeur obtenue en exécutant au préalable l'estimation *NB-QML* en profil tout en commençant par $\theta^{(-1)} = (2, 0.6, 0.3)'$ et $r^{(-1)} = 3$. Les valeurs de départ inobservables X_0 et $\lambda_0(\theta)$ de l'équation *INGARCH*(1, 1) sont estimées respectivement par

$$\tilde{X}_0 = \bar{X} \text{ et } \tilde{\lambda}_0(\theta) = \frac{\omega + \alpha \bar{X}}{1 - \beta} \simeq E(\lambda_t(\theta)), \text{ pour } \theta = (\omega, \alpha, \beta)' \in \Theta, \quad (5.2.5)$$

où \bar{X} est la moyenne empirique de la série X_1, \dots, X_n . Concernant l'Algorithme 4.5.1, qui n'est appliqué que dans le cas du modèle $NB_2-INGARCH$ (1.3.5), nous avons besoin d'estimer le paramètre de dispersion initial r^* . Pour cela nous mimons l'hypothèse GLM binomiale négative₂ (5.2.4), en prenant r^* comme solution de l'équation,

$$S^2 = \bar{X} \left(1 + \frac{1}{r^*} \bar{X}\right),$$

ç-à-d.

$$r^* = \frac{(\bar{X})^2}{S^2 - \bar{X}}, \quad (5.2.6)$$

où S^2 est la variance échantillonnale de X_1, \dots, X_n . Bien entendu, il n'y a pas de justification théorique pour ce choix. Nous avons tout simplement remplacé dans (5.2.4) la moyenne et la variance conditionnelles par leurs homologues inconditionnels. Pour ce choix, la série X_1, \dots, X_n doit être surdispersée (c'est-à-dire $S^2 > \bar{X}$), sinon r^* serait négatif, ce qui n'est pas valide.

La moyenne empirique des estimations, leurs écarts types (StD) empiriques, leurs erreurs standards asymptotiques (ASE) ainsi que les racines carrées empiriques de leurs erreurs quadratiques minimales sur les 500 répétitions sont rapportés dans les tableaux 5.2.1-5.2.3.

Le $RMSE$ d'une estimation $\hat{\theta}$ de θ_0 est calculé à partir de la formule

$$RMSE = \sqrt{bias^2 + StD^2},$$

où le $bias$ est la moyenne échantillonnale de $\hat{\theta} - \theta_0$ sur les 500 réplifications. Notons que les ASE sont obtenues à partir des variances asymptotiques des estimateurs $NB-QML$, données par les théorèmes 4.3.2 et 4.6.1, et du $P-QMLE$ (cf. Ahmad et Francq, 2016 et théorème 2.3.2).

θ_0		$\widehat{\theta}_P$	$\widehat{\theta}_1$	$\widehat{\theta}_4$
$\omega = 2$	<i>Moyenne</i>	1.9983	1.8804	1.9931
	<i>StD</i>	0.4067	0.3991	0.4094
	<i>ASE</i>	0.3979	0.4017	0.4041
	<i>RMSE</i>	0.4067	0.4166	0.4094
$\alpha_0 = 0.6$	<i>Moyenne</i>	0.6006	0.6147	0.6016
	<i>StD</i>	0.0363	0.0373	0.0377
	<i>ASE</i>	0.0360	0.0401	0.0398
	<i>RMSE</i>	0.0363	0.0400	0.0377
$\beta_0 = 0.3$	<i>Moyenne</i>	0.2982	0.2901	0.2977
	<i>StD</i>	0.0260	0.0266	0.0268
	<i>ASE</i>	0.0278	0.0280	0.0275
	<i>RMSE</i>	0.0260	0.0266	0.0268

Tableau 5.2.1. Moyenne, écart-type, erreur standard asymptotique et RMSE empirique de $\widehat{\theta}_r$ ($r = 1, 4$) et $\widehat{\theta}_P$ pour la série *INGARCH*(1, 1) poissonienne avec $\theta_0 = (2, 0.6, 0.3)'$ et $n = 1000$.

θ_0		$\widehat{\theta}_P$	$\widehat{\theta}_1$	$\widehat{\theta}_4$
$\omega = 2$	<i>Moyenne</i>	1.9375	2.0008	2.0516
	<i>StD</i>	0.6018	0.4465	0.4519
	<i>ASE</i>	0.5501	0.4565	0.4619
	<i>RMSE</i>	0.6050	0.4465	0.4548
$\alpha_0 = 0.6$	<i>Moyenne</i>	0.5721	0.5967	0.5844
	<i>StD</i>	0.0777	0.0459	0.0462
	<i>ASE</i>	0.0687	0.0470	0.0502
	<i>RMSE</i>	0.0825	0.0460	0.0488
$\beta_0 = 0.3$	<i>Moyenne</i>	0.3218	0.3031	0.3141
	<i>StD</i>	0.0744	0.0370	0.0376
	<i>ASE</i>	0.0802	0.0382	0.0388
	<i>RMSE</i>	0.0775	0.0371	0.0401

Tableau 5.2.2. Moyenne, écart-type, erreur standard asymptotique et RMSE empirique de $\widehat{\theta}_r$ ($r = 1, 4$) et $\widehat{\theta}_P$ pour la série *INGARCH*(1, 1) géométrique avec $\theta_0 = (2, 0.6, 0.3)'$ et $n = 1000$.

θ_0		$\hat{\theta}_P$	$\hat{\theta}_1$	$\hat{\theta}_4$	$\hat{\theta}_{\hat{r}}$
$\omega = 2$	<i>Moyenne</i>	2.1088	2.0711	2.0702	2.0316
	<i>Std</i>	0.4793	0.4658	0.4558	0.4508
	<i>ASE</i>	0.4981	0.4702	0.4664	0.4601
	<i>RMSE</i>	0.4914	0.4711	0.4611	0.4519
$\alpha_0 = 0.6$	<i>Moyenne</i>	0.6174	0.6109	0.6119	0.6166
	<i>Std</i>	0.0443	0.0491	0.0430	0.0427
	<i>ASE</i>	0.0533	0.0483	0.0446	0.0435
	<i>RMSE</i>	0.0475	0.0502	0.0446	0.0458
$\beta_0 = 0.3$	<i>Moyenne</i>	0.2712	0.2807	0.2796	0.2870
	<i>Std</i>	0.0305	0.0481	0.0296	0.0232
	<i>ASE</i>	0.0370	0.0450	0.0302	0.0262
	<i>RMSE</i>	0.0419	0.0518	0.0359	0.0265
$r_0 = 3$	<i>Moyenne</i>				2.9995
	<i>Std</i>	-	-	-	0.0330
	<i>RMSE</i>				0.0331

Tableau 5.2.3. Moyenne, écart-type, erreur standard asymptotique et RMSE empirique de $\hat{\theta}_r$ ($r = 1, 4$), $\hat{\theta}_P$ et $\hat{\theta}_{\hat{r}}$ pour la série $NB_2-INGARCH(1, 1)$ avec $r_0 = 3$, $\theta_0 = (2, 0.6, 0.3)'$ et $n = 1000$.

D'après les tableaux 5.2.1-5.2.3, notre analyse de Monte Carlo révèle que les paramètres sont largement bien estimés par toutes les méthodes consultées et que les résultats sont cohérents avec la théorie asymptotique. En effet, lorsque la distribution conditionnelle du modèle $INGARCH(1, 1)$ suit une distribution donnée, l'estimation QML calculée sur cette distribution est la meilleure par rapport aux autres en ce qui concerne son $RMSE$ qui est le plus petit. Plus précisément, dans le cas $INGARCH(1, 1)$ poissonien (cf. tableau 5.2.1), le $P-QMLE$ surpasse le $G-QMLE$ et le $pNB-QMLE$. De même, pour le modèle $INGARCH(1, 1)$ géométrique (cf. tableau 5.2.2), le $G-QMLE$ a un $RMSE$ plus petit que le $P-QMLE$ et le $NB-QMLE$ en profil, $\hat{\theta}_4$. Enfin, pour le modèle $NB_2-INGARCH(1, 1)$ avec comme paramètre de dispersion $r_0 = 3$ (cf. tableau 5.2.3), l'estimateur $\hat{\theta}_{\hat{r}}$ en quatre étapes surpasse le $QMLE$ poissonien, le $QMLE$ géométrique et le $QMLE-NB$ en profil, $\hat{\theta}_4$.

Comme nous l'avons mentionné précédemment, pour le cas $r = r_0$ le $QMLE$ se réduit à l'estimateur du maximum de vraisemblance et la méthode en deux étapes, $2SNB-MLE$, fournit toujours les meilleurs résultats. Dans ce qui suit, le cas $r = r_0 = 3$ est considéré où nous avons sélectionné $\theta^{(-1)} = (0.8, 0.8, 0.1)'$ comme valeur initiale.

θ_0		$\hat{\theta}_P$	$\hat{\theta}_1$	$\hat{\theta}_3$
$\omega = 0.8$	<i>Moyenne</i>	0.86	0.855	0.8584
	<i>StD</i>	0.1441	0.1743	0.1658
	<i>ASE</i>	0.1158	0.1325	0.1247
	<i>RMSE</i>	0.1561	0.1828	0.1757
$\alpha_0 = 0.8$	<i>Moyenne</i>	0.795	0.8042	0.803
	<i>StD</i>	0.0177	0.0256	0.0231
	<i>ASE</i>	0.0312	0.0368	0.0342
	<i>RMSE</i>	0.0184	0.0259	0.0233
$\beta_0 = 0.1$	<i>Moyenne</i>	0.1003	0.0935	0.0938
	<i>StD</i>	0.0206	0.034	0.0296
	<i>ASE</i>	0.0353	0.041	0.0382
	<i>RMSE</i>	0.0206	0.0347	0.0302

Tableau 5.2.4. Moyenne, écart-type, erreur standard asymptotique et RMSE empirique de $\hat{\theta}_r$ ($r = 1, 3$) et $\hat{\theta}_P$ pour la série *INGARCH*(1, 1) poissonienne avec $\theta_0 = (0.8, 0.8, 0.1)'$ et $n = 1000$.

θ_0		$\widehat{\theta}_P$	$\widehat{\theta}_1$	$\widehat{\theta}_3$
$\omega = 0.8$	<i>Moyenne</i>	1.0322	0.9515	0.9837
	<i>StD</i>	0.5234	0.4406	0.5174
	<i>ASE</i>	0.1717	0.1006	0.1031
	<i>RMSE</i>	0.5727	0.4659	0.549
$\alpha_0 = 0.8$	<i>Moyenne</i>	0.712	0.7456	0.7436
	<i>StD</i>	0.1088	0.108	0.1085
	<i>ASE</i>	0.1172	0.0542	0.055
	<i>RMSE</i>	0.1399	0.1209	0.1223
$\beta_0 = 0.1$	<i>Moyenne</i>	0.1118	0.1114	0.1111
	<i>StD</i>	0.0803	0.0637	0.0657
	<i>ASE</i>	0.0606	0.0352	0.0355
	<i>RMSE</i>	0.0812	0.0647	0.0666

Tableau 5.2.5. Moyenne, écart-type, erreur standard asymptotique et RMSE empirique de $\widehat{\theta}_r$ ($r = 1, 3$) et $\widehat{\theta}_P$ pour la série *INGARCH*(1, 1) géométrique avec $\theta_0 = (0.8, 0.8, 0.1)'$ et $n = 1000$.

θ_0		$\hat{\theta}_P$	$\hat{\theta}_1$	$\hat{\theta}_3$	$\hat{\theta}_{\hat{r}}$
$\omega = 0.8$	<i>Moyenne</i>	0.8758	0.8466	0.8422	0.8421
	<i>StD</i>	0.0581	0.0234	0.0235	0.0234
	<i>ASE</i>	0.1534	0.1355	0.0962	0.0962
	<i>RMSE</i>	0.0954	0.0522	0.0483	0.0482
$\alpha_0 = 0.8$	<i>Moyenne</i>	0.8314	0.7804	0.7883	0.7879
	<i>StD</i>	0.0320	0.0448	0.0367	0.0364
	<i>ASE</i>	0.0654	0.0577	0.038	0.038
	<i>RMSE</i>	0.0448	0.0489	0.0385	0.0383
$\beta_0 = 0.1$	<i>Moyenne</i>	0.0679	0.1054	0.1016	0.1019
	<i>StD</i>	0.0153	0.0273	0.0226	0.0223
	<i>ASE</i>	0.0498	0.0489	0.0328	0.0328
	<i>RMSE</i>	0.0356	0.0278	0.0227	0.0224
$r_0 = 3$	<i>Moyenne</i>				2.9345
	<i>StD</i>	-	-	-	0.2259
	<i>RMSE</i>				0.2352
\hat{r}_2	<i>Moyenne</i>				2.9366
	<i>StD</i>	-	-	-	0.2215
	<i>RMSE</i>				0.2304

Tableau 5.2.6. Moyenne, écart-type, erreur standard asymptotique et RMSE empirique de $\hat{\theta}_r$ ($r = 1, 3$), $\hat{\theta}_P$, $\hat{\theta}_{\hat{r}}$ et \hat{r}_2 pour la série $NB_2-INGARCH(1, 1)$ avec $r_0 = 3$, $\theta_0 = (0.8, 0.8, 0.1)'$ et $n = 1000$.

5.3 Comparaison sous le modèle $INAR(1)$

5.3.1 Comparaison vis-à-vis de l'efficacité asymptotique relative

Le modèle $INAR(1)$, (1.4.1), obéit à la fonction de lien GLM affine suivante

$$\begin{aligned} Var(X_t | \mathcal{F}_{t-1}) &= \alpha_0 (1 - \alpha_0) X_{t-1} + \sigma_0^2 \\ &= (1 - \alpha_0) E(X_t | \mathcal{F}_{t-1}) + \sigma_0^2 - (1 - \alpha_0) \omega_0. \end{aligned} \quad (5.3.1)$$

Notons que si $\frac{\sigma_0^2}{\omega_0} = 1 - \alpha_0 < 1$, auquel cas le terme d'innovation ε_t est *sous-dispersé*, alors la fonction de lien affine (5.3.1) se réduit à la condition du GLM poissonien (5.2.3) avec une constante de proportionnalité $1 - \alpha_0$. Par conséquent, le P - $QMLE$ serait asymptotiquement efficace dans la classe de tous les $QMLE$ appartenant à la famille exponentielle linéaire et donc il serait plus asymptotiquement efficace que le pNB - $QMLE$.

Cependant, si $\frac{\sigma_0^2}{\omega_0} \neq (1 - \alpha_0)$, alors aucun des deux estimateurs P - $QMLE$ et pNB - $QMLE$ n'est plus asymptotiquement efficace que l'autre.

De plus, dans tous les cas, le $2SNB$ - $QMLE$ est mal défini.

5.3.2 Comparaison en échantillons finis : étude de simulation

Dans le même but que la section 5.2.2, les résultats de simulations effectués pour une seule chaîne de valeurs de $\theta = (2, 0.9)'$ sont reportés dans les tableaux ci-dessous dont les séries simulées sont générées à partir des modèles *Poisson-INAR(1)* et *pNB₂-INAR(1)*.

θ_0		$\hat{\theta}_P$	$\hat{\theta}_1$	$\hat{\theta}_3$
$\omega = 2$	<i>Moyenne</i>	2.1342	2.1018	2.1326
	<i>StD</i>	0.2862	0.3122	0.3495
	<i>ASE</i>	0.3098	0.5709	0.4894
	<i>RMSE</i>	0.3161	0.3284	0.3738
$\alpha_0 = 0.9$	<i>Moyenne</i>	0.8932	0.8948	0.8932
	<i>StD</i>	0.014	0.0152	0.0173
	<i>ASE</i>	0.0155	0.0294	0.0251
	<i>RMSE</i>	0.0155	0.0161	0.0186

Tableau 5.3.1. Moyenne, écart-type, erreur standard asymptotique et RMSE empirique de $\hat{\theta}_r$ ($r = 1, 3$) et $\hat{\theta}_P$ pour le modèle *INAR(1)* avec innovation poissonnienne avec $\theta_0 = (2, 0.9)'$ et $n = 1000$.

θ_0		$\hat{\theta}_P$	$\hat{\theta}_1$	$\hat{\theta}_3$
$\omega = 2$	<i>Moyenne</i>	2.108	2.0887	2.1006
	<i>StD</i>	0.3071	0.3398	0.3674
	<i>ASE</i>	0.3207	0.5917	0.506
	<i>RMSE</i>	0.3255	0.3512	0.3809
$\alpha_0 = 0.9$	<i>Moyenne</i>	0.8946	0.8956	0.8949
	<i>StD</i>	0.0151	0.0169	0.0183
	<i>ASE</i>	0.0159	0.0307	0.026
	<i>RMSE</i>	0.016	0.0175	0.019

Tableau 5.3.2. Moyenne, écart-type, erreur standard asymptotique et RMSE empirique de $\hat{\theta}_r$ ($r = 1, 3$) et $\hat{\theta}_P$ pour le modèle $pNB_2-INAR(1)$ avec $r_0 = 3$, $\theta_0 = (2, 0.9)'$ et $n = 1000$.

Les tableaux 5.3.1- 5.3.2 montrent le fait qu'on ne peut pas conclure à propos de l'efficacité des estimateurs . Plus précisément, aucun des estimateurs $\hat{\theta}_P$ et $\hat{\theta}_3$ n'est pratiquement supérieur à l'autre. Par exemple, lorsque l'innovation du modèle $INAR(1)$ suit une distribution binomiale négative₂, l'estimateur QML calculé sur cette distribution n'est pas forcément le meilleur par rapport à la distribution poissonnienne.

Chapitre 6

Applications réelles

6.1 Introduction

A titre d'illustration, nous proposons d'appliquer la *NB-QMLE* en deux étapes donnée par l'Algorithme 4.5.1 à deux célèbres séries temporelles à valeurs entières au vu du modèle *INGARCH*(1,1) binomial négatif₂. La première est celle des données Polio (Zeger, 1988), tandis que la seconde est celle des données transaction (Fokianos et al, 2009). Le choix du modèle *NB₂-INGARCH*(1,1) est motivé par la surdispersion des séries citées. De plus, ces deux séries réelles ont été considérées par Zhu (2011) et Christou et Fokianos (2014) respectivement en utilisant ce modèle, mais ce via des méthodes d'estimation différentes. Cela nous permet de comparer leurs méthodes avec notre *2SNB-QMLE* proposée. Toutes les procédures ont été appliquées sur un ordinateur personnel en utilisant le logiciel *R*. Ainsi, l'optimisation (4.2.1) a été effectuée en utilisant la fonction `constrOptim()` de *R*.

6.2 Application à la série "Polio"

Le premier ensemble de données est le nombre mensuel de cas de poliomyélite aux États-Unis sur la période allant de 1970 à 1983 avec un total de $n = 168$ observations (voir figure 1.2.2). Cette série a été modélisée à l'origine par Zeger (1988) et utilisée plus tard par de nombreux auteurs (voir Zeger et Qaqish, 1988, Davis et al, 1999, Benjamin et al, 2003, Heinen, 2003, Davis et Wu, 2009, Zhu, 2011 parmi autres).

Zhu (2011) a ajusté un modèle $NB_2-INGARCH(1, 1)$ de la forme (1.3.6) à la série Polio. Comme précisé précédemment, ce modèle est légèrement différent du modèle (1.3.5). Tout d'abord, le paramètre de dispersion dans (1.3.6) est considéré comme un entier positif, ce qui est un peu restrictif. Deuxièmement, le paramètre de probabilité est $\frac{1}{1+\mu_t}$ plutôt que $\frac{r_0}{r_0+\lambda_t}$ dans (1.3.5). Ainsi, la moyenne conditionnelle du modèle (1.3.6) n'est pas sous la forme (1.6.1). Cependant, en prenant $\lambda_t = \frac{\mu_t}{r_0}$ nous trouvons le modèle (1.3.5) avec une paramétrisation différente. Zhu (2011) a estimé le modèle (1.3.6) à l'aide d'une estimation approximative du maximum de vraisemblance. Cette estimation consiste à maximiser la vraisemblance binomiale négative sur θ pour r fixé et ensuite choisir θ avec la plus grande vraisemblance sur toutes les valeurs choisies de $r \in \{1, \dots, \bar{r}\}$, pour un certain nombre entier positif fixé \bar{r} . Le modèle estimé de Zhu (2011) est donné par

$$\begin{aligned}
 X_t \mid \mathcal{F}_{t-1} &\sim \mathcal{NB} \left(\hat{r}, \frac{1}{1+\hat{\mu}_t} \right), & (6.2.1) \\
 \hat{r} &= 2, \\
 \begin{cases} \hat{\mu}_t = 0.31190 + 0.1843X_{t-1} + 0.1815\hat{\mu}_{t-1}, & 2 \leq t \leq 168 \\ \hat{\mu}_1 = \bar{X}, \end{cases}
 \end{aligned}$$

à partir duquel l'estimation de $E(X_t)$ est

$$2 \times \frac{0.3119}{1 - (0.1843 + 0.1815)} = 0.9836,$$

et le paramètre de persistance (ou de stabilité) est $0.1843 + 0.1815 = 0.3658$.

Ainsi, en suivant Christou et Fokianos (2014), l'application du P - $QMLE$ à la série Polio donne les résultats suivants.

<i>paramètres du modèle</i> <i>NB₂-INGARCH</i>	<i>Estimateurs</i>	<i>ASE</i>
ω_0	0.6401	0.1670
α_0	0.3501	0.1289
β_0	0.1821	0.0673

Tableau 6.2.1 : Estimateurs du $PQML$ et leurs erreurs standards asymptotiques pour le modèle $NB_2-INGARCH(1, 1)$ à partir de la série Polio.

Le modèle ajusté est alors donné par

$$X_t | \mathcal{F}_{t-1} \sim \mathcal{P}(\widehat{\lambda}_t^p), \tag{6.2.2}$$

$$\begin{cases} \widehat{\lambda}_t^p = 0.6401 + 0.3501X_{t-1} + 0.1821\widehat{\lambda}_{t-1}^p, & 2 \leq t \leq 168 \\ \widehat{\lambda}_1^p = \bar{X} = 1.3333, \end{cases}$$

avec comme paramètre de persistance 0.5322 et une moyenne estimée égale à

$$\frac{0.6401}{1 - (0.3501 + 0.1821)} = 1.3683.$$

Pour comparer avec l'ajustement de Zhu (2011) et le P - $QMLE$ de Christou et Fokianos (2014), nous avons estimé un modèle $NB_2-INGARCH(1, 1)$ (1.3.5) en utilisant la $2SNB-QMLE$ (Algorithme 4.5.1). Dans l'implémentation de l'Algorithme 4.5.1, nous avons utilisé les mêmes dispositifs que dans la Section 5.2.2. Plus précisément, le paramètre de dispersion initial r^* est calculé en utilisant (5.2.6) comme suit

$$r^* = \frac{(1.3333)^2}{3.5050 - 1.3333} = 0.8186,$$

tandis que les valeurs de départ de l'équation $INGARCH(1, 1)$ (1.3.1b) sont prises comme dans (5.2.5). Le paramètre initial $\theta^{(0)}$ de la moyenne conditionnelle du problème d'optimisation (4.2.1) est obtenu en exécutant au préalable le $QMLE$ géométrique sur la série Polio

en considérant comme paramètre initial $(2, 0.3, 0.6)'$. Les paramètres estimés du modèle et leurs erreurs standards asymptotiques (*ASE*) sont résumés dans le tableau 6.2.2. Ces dernières sont calculées à partir de la distribution asymptotique du *2SNB-QMLE* donnée par le théorème 4.6.1. En particulier, l'*ASE* de $\hat{\gamma}_2 = (\hat{r}_2)^{-1}$ est calculée à partir de (4.6.1b) et (4.6.3) alors que l'*ASE* de $\hat{\theta}_{\hat{r}_2}$ est obtenue à partir de (4.6.1d) et (4.6.2). Notons que l'*ASE* de \hat{r}_2 n'est pas disponible puisque la distribution de \hat{r}_2 n'a pas une forme usuelle, mais peut être simulée en utilisant une paramétrisation bootstrap.

<i>paramètres du modèle</i> <i>NB₂-INGARCH</i>	<i>Estimateurs :</i> $(\hat{\theta}_{\hat{r}_1}, \hat{\gamma}_2, \hat{r}_2)$	<i>ASE de</i> $\hat{\theta}_{\hat{r}_1}, \hat{\gamma}_2$
ω_0	0.6564	0.2050
α_0	0.3743	0.1580
β_0	0.1511	0.0935
$\gamma_0 = \frac{1}{r_0}$	0.3843	0.1945
r_0	2.6023	—

Tableau 6.2.2 : Estimateurs de la *2S-NBQML* et leurs erreurs standards asymptotiques pour le modèle *NB₂-INGARCH*(1, 1) à partir de la série Polio.

En utilisant la *2SNB-QMLE*, le modèle ajusté (1.3.5) est donné par

$$X_t | \mathcal{F}_{t-1} \sim \mathcal{NB} \left(\hat{r}_2, \frac{\hat{r}_2}{\hat{r}_2 + \hat{\lambda}_t} \right), \quad (6.2.3)$$

$$\hat{r}_2 = 2.6023,$$

$$\begin{cases} \hat{\lambda}_t = 0.6564 + 0.3743X_{t-1} + 0.1511\hat{\lambda}_{t-1}, & 2 \leq t \leq 168 \\ \hat{\lambda}_1 = \bar{X} = 1.3333, \end{cases}$$

Avec paramètre de persistance $0.3743 + 0.1511 = 0.5254$. Notons que notre estimation de la moyenne $E(X_t)$ est telle que

$$\frac{0.6564}{1 - (0.3743 + 0.1511)} = 1.3834,$$

qui est, en effet, plus proche à la moyenne échantillonnale $\bar{X} = 1.3333$ que la moyenne estimée, 0.9836, donnée par le modèle de Zhu (2011). Néanmoins, la moyenne estimée donnée par le P - $QMLE$ de Christou et Fokianos (2014) est légèrement meilleur que la moyenne estimée induite par la $2SNB$ - $QMLE$.

D'autre part, certaines propriétés des résidus sont exhibées dans la figure 6.2.1. En effet, à partir des fonctions d'autocorrélation et d'autocorrélation partielle de la figure 6.2.1 (panneaux (a) et (b)), les résidus ressemblent à un bruit blanc. Cependant, une inspection visuelle (voir la figure 6.2.1, panneaux (c) et (d)) révèle que l'hypothèse de normalité des résidus est intenable. En résumé, en ce qui concerne la stabilité du modèle estimé, la signification de ses coefficients et l'analyse résiduelle dans la figure 6.2.1, on peut conclure que le modèle estimé est acceptable.

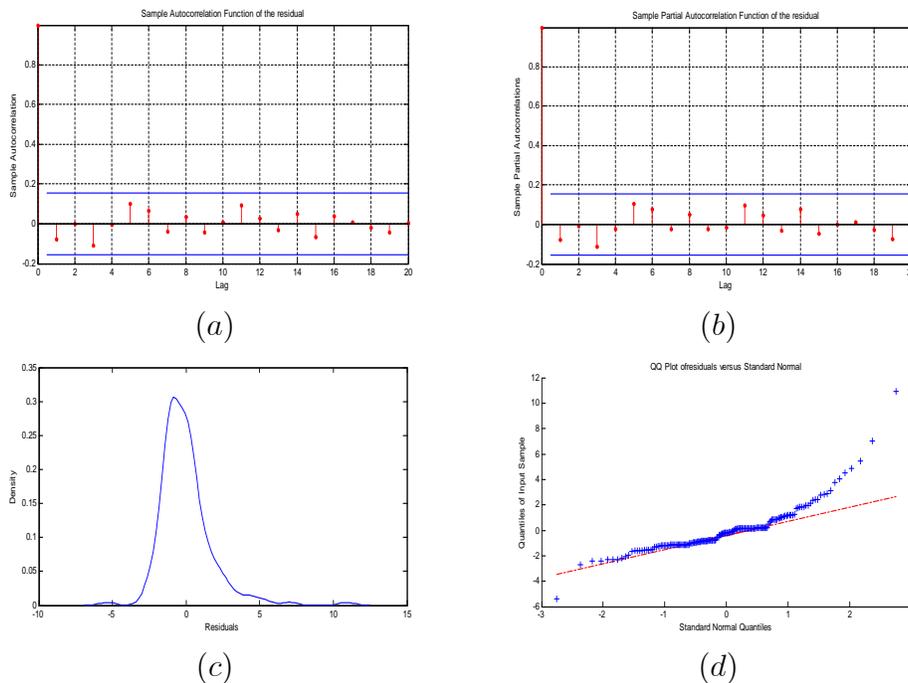


Figure 6.2.1 : Analyse résiduelle pour la série Polio.

- (a) autocorrélation échantillonnale des résidus.
- (b) autocorrélation partielle échantillonnale des résidus.
- (c) densité de noyau des résidus.
- (d) QQ-plot des résidus par rapport à la distribution normale.

Maintenant, nous comparons les performances des échantillons de notre ajustement avec celles de Zhu (2011) et le P - $QMLE$ de Christou et Fokianos (2014). Le tableau 6.2.3 fournit la somme des carrés résiduelle (RSS) induite par les modèles (6.2.1), (6.2.2) et (6.2.3).

Ces RSS sont donnés respectivement par

$$RSS(\widehat{\lambda}_t) = \sum_{t=2}^{168} (X_t - \widehat{\lambda}_t)^2,$$

$$RSS(2\widehat{\mu}_t) = \sum_{t=2}^{168} (X_t - 2\widehat{\mu}_t)^2,$$

$$RSS(\widehat{\lambda}_t^p) = \sum_{t=2}^{168} (X_t - \widehat{\lambda}_t^p)^2.$$

en considérant comme valeurs initiales $\widehat{\lambda}_1 = \widehat{\lambda}_1^p = \widehat{\mu}_1 = \bar{X}$. Cette dernière valeur initiale a été considérée par Zhu (2011).

Prédicteurs	$\widehat{\lambda}_t$	$2\widehat{\mu}_t$	$\widehat{\lambda}_t^p$
RSS	535.1793	540.6634	533.5275

Tableau 6.2.3 : Somme des carrés résiduelle (RSS) des prédicteurs $\widehat{\lambda}_t$ (6.2.3), $2\widehat{\mu}_t$ (6.2.1) et $\widehat{\lambda}_t^p$ (6.2.2) pour la série Polio.

Du tableau 6.2.3, on constate que notre modèle estimé par la $2SNB$ - $QMLE$ (Algorithme 4.5.1) dépasse légèrement le modèle de Zhu (2011) avec une somme des carrés résiduelle plus petite (RSS). Comme la moyenne conditionnelle peut être influencée par le choix des valeurs initiales, nous avons calculé plusieurs RSS correspondants aux modèles (6.2.1), (6.2.2) et (6.2.3) à partir de plusieurs valeurs initiales $\widehat{\lambda}_1$, $\widehat{\lambda}_1^p$ et $\widehat{\mu}_1$; Les résultats non signalés étaient pratiquement les mêmes.

Enfin, la figure 6.2.2 exhibe les données Polio avec sa moyenne conditionnelle estimée $\widehat{\lambda}_t$ et sa variance conditionnelle estimée donnée par

$$\widehat{v}_t = \widehat{\lambda}_t \left(1 + \frac{1}{\widehat{r}_2} \widehat{\lambda}_t \right),$$

où le phénomène de surdispersion semble reproduit.

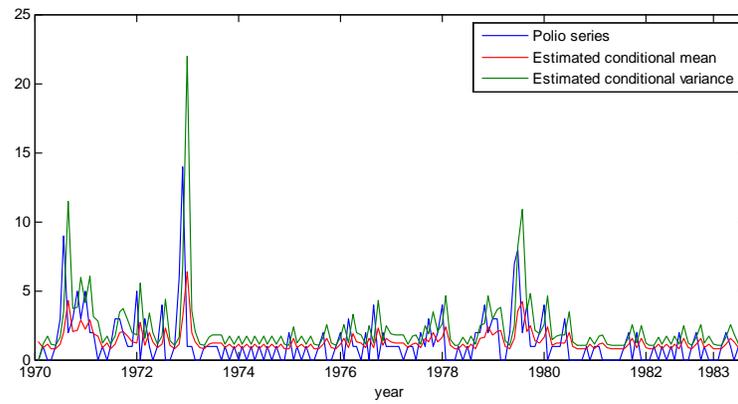


Figure 6.2.2 : Série Polio et leur moyenne et variance conditionnelles estimées ($2SNB-QML$).

6.3 Application à la série "transaction"

Le deuxième ensemble de données est le nombre de transactions par minute pour le stock Ericsson B au cours de 05 Juillet, 2002. Cette série a un total de $n = 460$ observations représentant la transaction d'environ 8 heures (de 09 :35 à 17 :14, cf. Figure 1.2.1). Elle a été utilisée par Fokianos et *al* (2009), Davis et Liu (2009) et Christou et Fokianos (2014) entre autres.

En utilisant le *QMLE* poissonien, Christou et Fokianos (2014) ont adapté un modèle $NB_2-INGARCH(1, 1)$ (1.3.5) aux données de transaction. Ils ont trouvé la spécification suivante

$$X_t | \mathcal{F}_{t-1} \sim \mathcal{NB} \left(\hat{r}, \frac{\hat{r}}{\hat{r} + \hat{\mu}_t} \right), \quad (6.3.1)$$

$$\hat{r} = 7.0220,$$

$$\begin{cases} \hat{\mu}_t = 0.5808 + 0.1986X_{t-1} + 0.7445\hat{\mu}_{t-1}, & 2 \leq t \leq 460 \\ \hat{\mu}_1 = 0, \end{cases}$$

avec une forte valeur du paramètre de persistance 0.9431 et une moyenne estimée à

$$\frac{0.5808}{1 - 0.9431} = 10.2070.$$

Motivé par le fait que l'estimateur $2SNB-QML$ (Algorithme 4.5.1) est plus asymptotiquement efficace que le $P-QMLE$ dans le contexte du modèle $NB_2-INGARCH$ (cf. section 5.2.1), nous avons appliqué l'estimation précédente à la série transaction en gardant les mêmes dispositifs utilisés pour les données Polio. En effet, à partir de (5.2.6), le paramètre de dispersion initial r^* est considéré comme

$$r^* = \frac{(9.8239)^2}{23.7532 - 9.8239} = 6.9285,$$

tandis que les valeurs de départ de l'équation $INGARCH(1, 1)$ (1.3.1b) sont définies conformément à (5.2.5). Les estimations des paramètres ainsi que leurs erreurs standards asymp-

totiques (*ASE*) sont résumées dans le tableau 6.3.1.

<i>paramètres du modèle</i> <i>NB₂-INGARCH</i>	<i>Estimateurs</i> $(\hat{\theta}_{\hat{r}_1}, \hat{\gamma}_2, \hat{r}_2)$	<i>ASE de</i> $\hat{\theta}_{\hat{r}_1}, \hat{\gamma}_2$
ω_0	0.7996	0.4034
α_0	0.7928	0.0650
β_0	0.1249	0.0340
$\gamma_0 = \frac{1}{r_0}$	0.1279	0.0241
r_0	7.8199	—

Tableau 6.3.1 : estimateurs de la *2S-NBQML* et leurs erreurs standards asymptotiques pour le modèle *NB₂-INGARCH(1, 1)* à partir de la série transaction.

Ainsi notre modèle ajusté *NB₂-INGARCH(1, 1)* à partir de la série transaction en utilisant la *2SNB-QMLE* est donné par

$$X_t | \mathcal{F}_{t-1} \sim \mathcal{NB} \left(\hat{r}_2, \frac{\hat{r}_2}{\hat{r}_2 + \hat{\lambda}_t} \right), \quad (6.3.2)$$

$$\hat{r}_2 = 7.8199,$$

$$\begin{cases} \hat{\lambda}_t = 0.7996 + 0.7928X_{t-1} + 0.1249\hat{\lambda}_{t-1}, & 2 \leq t \leq 460 \\ \hat{\lambda}_1 = \bar{X} = 9.8134, \end{cases}$$

dont la valeur estimée du paramètre de persistance vaut 0.9177 et celle de la moyenne égale à

$$\frac{0.7996}{1 - 0.9177} = 9.7157$$

qui est plus proche de la moyenne échantillonnale $\bar{X} = 9.8239$ que la moyenne estimée à partir de la spécification de Christou et Fokianos(2014).

La figure 6.3.1 présente la fonction d'autocorrélation de l'échantillon (panneau (a)), la fonction d'autocorrélation partielle de l'échantillon (panneau (b)), la densité de noyau (panneau (c)) et le QQ-plot (panneau (d)) des résidus du modèle (6.3.2). Il s'avère que l'hypothèse

selon laquelle les résidus forment un bruit blanc non gaussien est fortement soutenable.

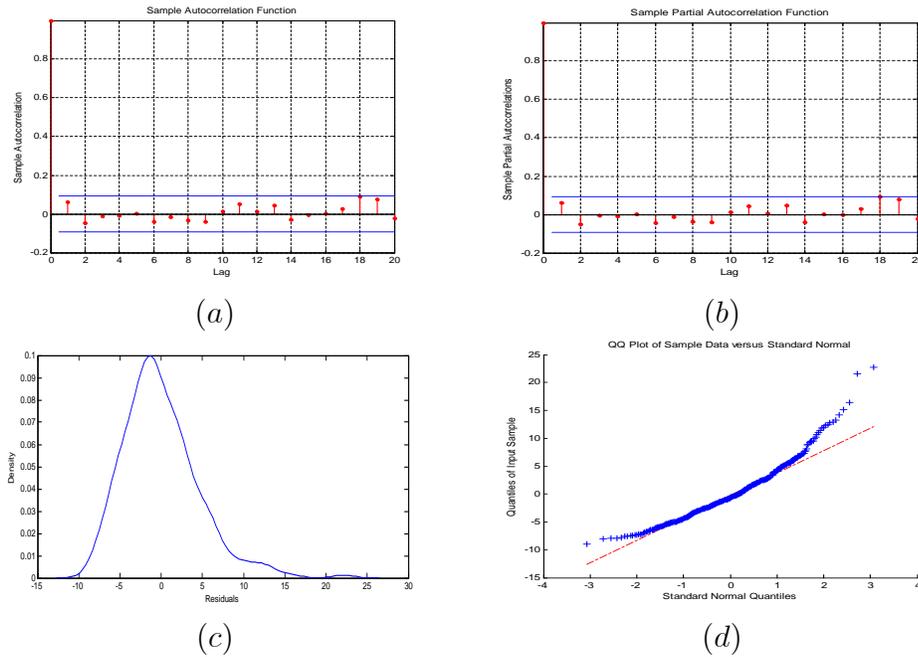


Figure 6.3.1 : Analyse résiduelle pour la série transaction.

- (a) autocorrélation échantillonnale des résidus.
- (b) autocorrélation partielle échantillonnale des résidus.
- (c) densité de noyau des résidus.
- (d) QQ-plot des résidus par rapport à la distribution normale.

Ensuite, nous comparons le RSS de notre ajustement (6.3.2) à celui de Christou et Fokianos (2014) donné par (6.3.1). Ces RSS sont donnés respectivement par

$$RSS(\hat{\lambda}_t) = \sum_{t=2}^{460} (X_t - \hat{\lambda}_t)^2,$$

$$RSS(\hat{\mu}_t) = \sum_{t=2}^{460} (X_t - \hat{\mu}_t)^2,$$

en considérant comme valeurs initiales $\hat{\lambda}_1 = \hat{\mu}_1 = \bar{X}$.

En raison des fortes valeurs des paramètres de persistance dans les deux modèles, les RSS peuvent être influencés par les valeurs de départ de la taille modérée de l'échantillon de la série des transactions. Nous avons donc commencé les équations (6.3.1) et (6.3.2) à partir de plusieurs valeurs initiales (voir le tableau 6.3.2) bien que Christou et Fokianos (2014) ont pris $\hat{\mu}_1 = 0$.

Prédicteurs	$\hat{\lambda}_t$	$\hat{\mu}_t$	$\hat{\lambda}_t$	$\hat{\mu}_t$	$\hat{\lambda}_t$	$\hat{\mu}_t$
Valeurs initiales	0	0	9.8239	9.8239	10.2070	10.2070
$\hat{\lambda}_1, \hat{\mu}_1$						
RSS	10400.6733	10422.8003	9809.6645	9943.0150	9796.8644	9933.0780

Tableau 6.3.2 : Somme des carrés résiduelle (RSS) des prédicteurs $\hat{\lambda}_t$ (6.3.2) et $\hat{\mu}_t$ (6.3.1) pour la série des transactions.

Du tableau 6.3.2, on constate que le modèle (6.3.2) estimé par la méthode $2SNB-QML$ a le plus petit RSS pour toutes les valeurs initiales choisies. La figure 6.3.2 exhibe la série des transactions avec sa moyenne conditionnelle estimée $\hat{\lambda}_t$ et sa variance conditionnelle estimée donnée par

$$\hat{v}_t = \hat{\lambda}_t \left(1 + \frac{1}{\hat{r}_2} \hat{\lambda}_t \right),$$

où le phénomène de surdispersion est mis en évidence.

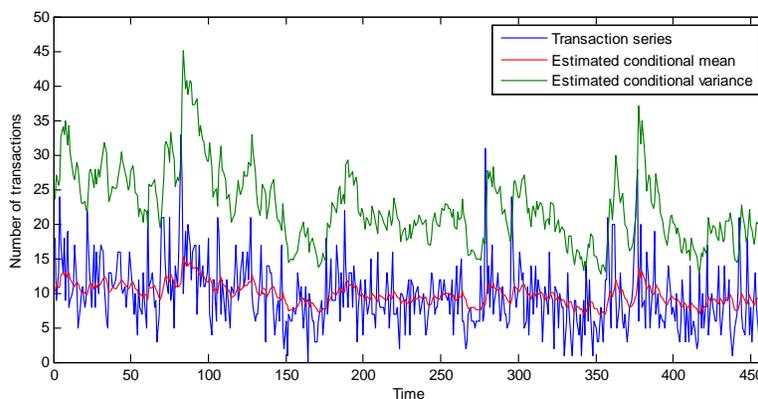


Figure 6.3.2 : La série des transactions et leur moyenne et variance conditionnelles estimées.

Chapitre 7

Conclusion générale

Dans cette thèse, nous avons proposé trois estimateurs du QML , à savoir l'estimateur du QML géométrique ($G-QMLE$), l'estimateur du QML binomial négatif en profil ($pNB-QMLE$) et l'estimateur du QML binomial négatif en deux étapes ($2SNB-QMLE$), et ce pour une classe générale de modèles de séries temporelles à valeurs entières. Ces estimateurs sont fortement consistants et asymptotiquement gaussiens sous des hypothèses générales et assez faibles mais sous la bonne spécification de la moyenne conditionnelle. En particulier, ils sont robustes devant une mauvaise spécification de la vraie distribution conditionnelle du modèle. De plus, sous la fonction de lien GLM binomiale négative de type 2, l'estimateur $NB-QMLE$ en deux étapes est plus efficace asymptotiquement que le Poisson $QMLE$ et est particulièrement bien adapté aux séries surdispersées. En outre, il est asymptotiquement efficace dans la classe de tous les estimateurs du QML appartenant à la famille exponentielle linéaire. Ainsi, le $NB-QMLE$ en deux étapes peut être considéré comme une bonne alternative à l'estimateur du maximum de vraisemblance (pour des modèles avec des distributions conditionnelles binomiales négatives de type 2), qui souffre de la non-robustesse à la spécification incorrecte de la vraie distribution conditionnelle et dont le calcul est très fastidieux.

Le travail que nous avons réalisé dans cette thèse pourrait être complété en étudiant la distribution asymptotique des estimateurs proposés quand le vrai paramètre se trouve sur le bord de l'espace paramétrique comme dans le cas du P-QMLE (Ahmad et Francq, 2016). D'autre part, des tests portemanteaux pour l'adéquation du modèle $INGAR$ peuvent être considérés en utilisant les estimateurs $NB-QMLEs$ proposés au lieu de l'estimateur du

P-*QML* (Christou et Neumann, 2013 ; Schweer, 2016). Enfin, des généralisations des *NB-QMLEs* proposés à des versions multivariées du modèle *INGAR* nous semblent attrayantes.

Notons enfin qu'en dépit de la généralité du modèle *INGAR*, il existe d'autres classes de modèles de séries chronologiques bien connus et auxquelles nous nous sommes pas intéressés dans le cadre de cette thèse. Nous citons en particulier l'une des classes les plus anciennes qu'est celle des modèles *DARMA* (discrète *ARMA* : Jacobs et Lewis, 1978a ; 1978b ; 1983), les modèles *GLARMA* (Generalized Linear *ARMA*, e.g. Dunsmuir, 2015) et les modèles espace d'état à valeurs entières (e.g. Brandt et *al*, 2000 ; Davis et Wu, 2009).

Chapitre 8

Annexe A

8.1 Glossaire

$p.s.$	Presque sûre (sûrement).
$\xrightarrow[n \rightarrow \infty]{p.s.}$	Convergence presque sûre quand $n \rightarrow \infty$.
$\xrightarrow[n \rightarrow \infty]{\mathcal{L}}$	Convergence en loi quand $n \rightarrow \infty$.
$\xrightarrow[n \rightarrow \infty]{p}$	Convergence en probabilité quand $n \rightarrow \infty$.
$o_p(1)$	Terme convergent en probabilité vers 0 quand $n \rightarrow \infty$.
$o_{p.s.}(1)$	Terme convergent presque sûrement vers 0 quand $n \rightarrow \infty$.
$O_p(1)$	Terme borné en probabilité quand $n \rightarrow \infty$.
\mathbb{N}	Ensemble des entiers naturels.
\mathbb{N}^*	Ensemble des nombres entiers positifs.
\mathbb{R}	Ensemble des nombres réels.
\mathbb{Z}	Ensemble des nombres entiers.
iid	Indépendant et identiquement distribué.
$X \sim \mathcal{P}(\mu)$	La variable aléatoire X suit une distribution poissonnienne.
$X \sim \mathcal{NB}(r, p)$	La variable aléatoire X suit une distribution binomiale négative.
$X \sim \mathcal{NB}\left(r\lambda^{2-K}, \frac{r\lambda^{2-K}}{r\lambda^{2-K} + \lambda}\right)$	La variable aléatoire X suit une distribution K -binomiale négative.
$X \sim \mathcal{NB}_1$	La variable aléatoire X suit une distribution binomiale négative de type 1.
$X \sim \mathcal{NB}_2$	La variable aléatoire X suit une distribution binomiale négative de type 2.
$X \sim \mathcal{G}(p)$	La variable aléatoire X suit une distribution géométrique.

<i>ARMA</i>	AutoRegressive Moving Average - Autorégressif Moyenne mobile.
<i>GARCH</i>	Generalized <i>ARCH</i> - Autorégressif conditionnellement hétéroscédastique généralisé.
<i>GARMA</i>	Generalized <i>ARMA</i> - <i>ARMA</i> généralisés.
<i>INAR</i>	IN-teger AutoregRession - <i>AR</i> entier.
<i>GINAR</i>	Generalized <i>INAR</i> - <i>INAR</i> généralisé.
<i>INGARCH</i>	Integer valued <i>GARCH</i> - Autorégressif conditionnellement hétéroscédastique généralisé à valeurs entières.
<i>NB_K-INGARCH</i>	Negative Binomial _K <i>INGARCH</i> - <i>INGARCH</i> <i>K</i> -binomial négatif.
<i>NB₁-INGARCH</i>	Negative Binomial ₁ <i>INGARCH</i> - <i>INGARCH</i> binomial négatif de type 1.
<i>NB₂-INGARCH</i>	Negative Binomial ₂ <i>INGARCH</i> - <i>INGARCH</i> binomial négatif de type 2.
<i>INGAR</i>	INteger Generalized AutoRegression - <i>AR</i> généralisé à valeurs entières.
<i>GLM</i>	Generalized Linear Models - Modèles Linéaires Généralisés.
<i>ML</i>	Maximum Likelihood - Maximum de vraisemblance.
<i>P-ML</i>	Poisson Maximum Likelihood - Maximum de vraisemblance poissonienne.
<i>G-ML</i>	Geometric Maximum Likelihood - Maximum de vraisemblance géométrique.
<i>NB-ML</i>	Negative binomial Maximum Likelihood - Maximum de vraisemblance binomiale négative.
<i>pNB-ML</i>	Profile negative binomial Maximum Likelihood - Maximum de vraisemblance binomiale négative en Profil.
<i>QML</i>	Quasi-Maximum Likelihood - Quasi-maximum de vraisemblance.
<i>P-QML</i>	Poisson <i>QML</i> - Quasi-maximum de vraisemblance poissonienne.
<i>G-QML</i>	Geometric <i>QML</i> - Quasi-maximum de vraisemblance géométrique.
<i>NB-QML</i>	Negative binomial <i>QML</i> - Quasi-maximum de vraisemblance binomiale négative.
<i>pNB-QML</i>	Profile negative binomial <i>QML</i> - Quasi-maximum de vraisemblance binomiale négative en Profil.
<i>MLE</i>	Maximum Likelihood Estimate - Estimateur du maximum de vraisemblance.
<i>P-MLE</i>	Poisson <i>MLE</i> - Estimateur du maximum de vraisemblance poissonien.
<i>G-MLE</i>	Geometric <i>MLE</i> - Estimateur du maximum de vraisemblance géométrique.
<i>NB-MLE</i>	Negative binomial <i>MLE</i> - Estimateur du maximum de vraisemblance binomial négatif.
<i>pNB-MLE</i>	Profile negative binomial <i>MLE</i> - Estimateur du maximum de vraisemblance binomial négatif en profil.

<i>QMLE</i>	Quasi-Maximum Likelihood Estimate - Estimateur du quasi-maximum de vraisemblance.
<i>P-QMLE</i>	Poisson <i>QMLE</i> - Estimateur du quasi-maximum de vraisemblance poissonien.
<i>G-QMLE</i>	Geometric <i>QMLE</i> - Estimateur du quasi-maximum de vraisemblance géométrique.
<i>NB-QMLE</i>	Negative binomial <i>QMLE</i> - Estimateur du quasi-maximum de vraisemblance binomial négatif.
<i>pNB-QMLE</i>	Profile negative binomial <i>QMLE</i> - Estimateur du quasi-maximum de vraisemblance binomial négatif en profil.
<i>2SNB-QMLE</i>	Two stage negative binomial <i>QMLE</i> - Estimateur du quasi-maximum de vraisemblance binomial négatif en deux étapes.
<i>RMSE</i>	Root Mean Square Error - Racine de l'erreur quadratique moyenne.
<i>StD</i>	Standard deviation - Ecart type.
<i>ASE</i>	Asymptotic standard errors - Erreurs standards asymptotiques.
<i>RSS</i>	Residual sum of squares - Somme des carrés résiduelle.

8.2 Liste des Tableaux

1.2.1. Données de la série Polio, $n = 168$ observations.....	2
1.2.2. Données de la série nombre de transactions, $n = 460$ observations.....	3
1.3.1 Différentes valeurs des paramètres $(\omega_0, \beta_{01}, \alpha_{01})$ et les moyennes correspondantes aux 4 processus $INGARCH(1, 1)$	9
5.2.1 Moyenne, écarts-types, erreur standard asymptotique et $RMSE$ empirique de $\hat{\theta}_r$ ($r = 1, 4$) et $\hat{\theta}_P$ pour la série $INGARCH(1, 1)$ poissonnienne avec $\theta_0 = (2, 0.6, 0.3)'$ et $n = 1000$	72
5.2.2 Moyenne, écart-type, erreur standard asymptotique et $RMSE$ empirique de $\hat{\theta}_r$ ($r = 1, 4$) et $\hat{\theta}_P$ pour la série $INGARCH(1, 1)$ géométrique avec $\theta_0 = (2, 0.6, 0.3)'$ et $n = 1000$	73
5.2.3 Moyenne, écart-type, erreur standard asymptotique et $RMSE$ empirique de $\hat{\theta}_r$ ($r = 1, 4$), $\hat{\theta}_P$ et $\hat{\theta}_{\hat{r}}$ pour la série $NB_2-INGARCH(1, 1)$ avec $r_0 = 3$, $\theta_0 = (2, 0.6, 0.3)'$ et $n = 1000$	74
5.2.4 Moyenne, écart-type, erreur standard asymptotique et $RMSE$ empirique de $\hat{\theta}_r$ ($r = 1, 3$) et $\hat{\theta}_P$ pour la série $INGARCH(1, 1)$ poissonnienne avec $\theta_0 = (0.8, 0.8, 0.1)'$ et $n = 1000$	76
5.2.5 Moyenne, écart-type, erreur standard asymptotique et $RMSE$ empirique de $\hat{\theta}_r$ ($r = 1, 3$) et $\hat{\theta}_P$ pour la série $INGARCH(1, 1)$ géométrique avec $\theta_0 = (0.8, 0.8, 0.1)'$ et $n = 1000$	77
5.2.6 Moyenne, écart-type, erreur standard asymptotique et $RMSE$ empirique de $\hat{\theta}_r$ ($r = 1, 3$), $\hat{\theta}_P$, $\hat{\theta}_{\hat{r}}$ et \hat{r}_2 pour la série $NB_2-INGARCH(1, 1)$ avec $r_0 = 3$, $\theta_0 = (0.8, 0.8, 0.1)'$ et $n = 1000$	78
5.3.1 Moyenne, écart-type, erreur standard asymptotique et $RMSE$ empirique de $\hat{\theta}_r$ ($r = 1, 3$) et $\hat{\theta}_P$ pour le modèle $INAR(1)$ avec innovation poissonnienne avec $\theta_0 = (2, 0.9)'$ et $n = 1000$	80

5.3.2 Moyenne, écart-type et *RMSE* de $\hat{\theta}_r$ ($r = 1, 3$) et $\hat{\theta}_P$ pour le modèle pNB_2 -*INAR*(1) avec $r_0 = 3$, $\theta_0 = (2, 0.9)'$ et $n = 1000$81

6.2.1 Estimateurs du *PQML* et leurs erreurs standards asymptotiques pour le modèle NB_2 -*INGARCH*(1, 1) à partir de la série Polio.....84

6.2.2 Estimateurs de la *2S-NBQML* et leurs erreurs standards asymptotiques pour le modèle NB_2 -*INGARCH*(1, 1) à partir de la série Polio.....85

6.2.3 Somme des carrés résiduelle (*RSS*) des prédicteurs $\hat{\lambda}_t$ (6.2.3), $2\hat{\mu}_t$ (6.2.1) et $\hat{\lambda}_t^p$ (6.2.2) pour la série Polio.....87

6.3.1 Estimateurs de la *2S-NBQML* et leurs erreurs standards asymptotiques pour le modèle NB_2 -*INGARCH*(1, 1) à partir de la série transaction.....90

6.3.2 Somme des carrés résiduelle (*RSS*) des prédicteurs $\hat{\lambda}_t$ (6.3.2) et $\hat{\mu}_t$ (6.3.1) pour la série des transactions.....92

8.3 Liste des Figures

1.2.1 Graphe de la série transaction.....4

1.2.2 Graphe de la série Polio.....4

1.2.3 Histogramme des deux séries..... .5
 (a) Série Polio. (b) Série transaction.

1.2.4 Corrélogramme de la série Polio..... .5

1.3.1 Histogramme du processus *INGARCH*(5, 2) simulé avec la fonction densité poissonnienne en rouge.....7

1.3.2 Histogrammes des processus *INGARCH*(1, 1) X_t, Y_t, Z_t et K_t respectivement.....9

1.3.3 200 observations et ses fonctions d'autocorrélations simples à partir du modèle (1.3.9) pour différentes valeurs des paramètres.....15
 (a) : $w_0 = 0.5, \alpha_{01} = 2/3, \beta_{01} = -0.5$.
 (b) : $w_0 = 0.5, \alpha_{01} = 1/3, \beta_{01} = 0.5$.

(c) : $w_0 = 0.5$, $\alpha_{01} = -1/4$, $\beta_{01} = -0.5$.

6.2.1	Analyse résiduelle pour la série Polio.....	86
	(a) autocorrélation échantillonnale des résidus.	
	(b) autocorrélation partielle échantillonnale des résidus.	
	(c) densité de noyau des résidus.	
	(d) QQ-plot des résidus par rapport à la distribution normale.	
6.2.2	Série Polio et leur moyenne et variance conditionnelles estimées (<i>2SNB-QML</i>).....	88
6.3.1	Analyse résiduelle pour la série transaction.	91
	(a) autocorrélation échantillonnale des résidus.	
	(b) autocorrélation partielle échantillonnale des résidus.	
	(c) densité de noyau des résidus.	
	(d) QQ-plot des résidus par rapport à la distribution normale.	
6.3.2	La série des transactions et leur moyenne et variance conditionnelles estimées.....	92

Bibliographie

- [1] Ahmad, A. et Francq, C. (2016). Poisson qmle of count time series models. *Journal of Time Series Analysis*, **37**, 291-314.
- [2] Aknouche, A. (2013). Sur l'analyse des séries chronologiques à valeurs entières : état de l'art. Prépublication.
- [3] Aknouche, A. (2014). *Modèles de séries chronologiques homogènes et périodiquement homogènes : stabilité, inférence et applications*. Mémoire d'Habilitation à Diriger des Recherches (HDR) - Mathématiques appliquées, Section 26, Université de Lille 3, France.
- [4] Aknouche, A. (2015). *Unified quasi-maximum likelihood estimation theory for stable and unstable Markov bilinear processes*. Book-Chapter in *Progress in economics research*, Nova Publishers, New York.
- [5] Aknouche, A., Al-Eid, E. et Demouche, N. (2018a). Generalized quasi-maximum likelihood inference for periodic conditionally heteroskedastic models. *Statistical Inference for Stochastic Processes*, forthcoming. DOI : 10.1007/s11203-017-9160-x.
- [6] Aknouche, A. et Bendjeddou, S. (2017). Estimateur du quasi-maximum de vraisemblance géométrique d'une classe générale de modèles de séries chronologiques à valeurs entières. *Comptes Rendus Mathématique*, **355**, 99-104.
- [7] Aknouche, A., Bendjeddou, S. et Touche, N. (2018b). Negative binomial quasi-likelihood inference for general integer-valued time series models. *Journal of Time Series Analysis*, forthcoming, DOI : 10.1111/jtsa.12277.
- [8] Al-Osh, M.A. et Aly, E-E A.A. (1992). First-order autoregressive time series with negative binomial and geometric marginals. *Commun. Statist- Theory Meth*, **21**, 2483-2492.

- [9] Al-Osh, M.A. et Alzaid, A.A. (1987). First-order integer-valued autoregressive ($INAR(1)$) process. *Journal of Time Series Analysis*, **8**, 261-275.
- [10] Al-Osh, M.A. et Alzaid, A.A. (1988). Integer-valued moving average ($INMA$) process. *Statist*, **29**, 281-300.
- [11] Al-Osh, M.A. et Alzaid, A.A. (1990). An integer-valued p th-order autoregressive structure ($INAR(P)$) process. *J. Appl. Probab.*, **27**(2), 314-324.
- [12] Al-Osh, M.A. et Alzaid, A.A. (1991). Binomial autoregressive moving average models. *Stoch. Models*, **7**, 261-282.
- [13] Al-Osh, M.A. et Alzaid, A.A. (1993). Some autoregressive moving average processes with generalized Poisson marginal distributions. *Ann. Inst. Statist. Math.*, **45**(2), 223-232.
- [14] Benjamin, M.A., Rigby, R.A. et Stasinopoulos, D.M. (2003). Generalized autoregressive moving average models. *Journal of the American Statistical Association*, **98**, 214-223.
- [15] Billingsley, P. (2008). *Probability and measure*, 3rd ed, John Wiley.
- [16] Bourguignon, M. (2016). Poisson-geometric $INAR(1)$ process for modeling count time series with overdispersion. *Statistica Neerlandica*, **70**, 176-192.
- [17] Brandt, Patrick T., John T. Williams, Benjamin O. Fordham et Brain Pollins. (2000). Dynamic Modeling for Persistent Event-Count Time Series. *American Journal of Political Science*, **44**, 823-43.
- [18] Brumback, B. A., Ryan, L. M., Schwartz, J. D., Neas, L. M., Stark, P. C., Burge, H. A. (2000). Transitional regression models with application to environmental time series. *Journal of the American Statistical Association*, **85**, 16-27.
- [19] Cameron, A.C. et Trivedi, P.K. (1986). Econometric models based on count data : Comparisons and applications of some estimators and tests. *Journal of Applied Econometrics*, **1**, 29-53.
- [20] Cameron, C. et Trivedi, P. (2013). *Regression analysis of count data*. Cambridge University Press, 2nd ed. New York.
- [21] Chen, C.W.S., So, M., Li, J.C. et Sriboonchitta, S. (2016). Autoregressive conditional negative binomial model applied to over-dispersed time series of counts. *Statistical Methodology*, **31**, 73-90.

- [22] Christou, V. (2013). *Statistical theory for mixed poisson time series models*. PhD thesis, University of Cyprus.
- [23] Christou, V. et Fokianos, K. (2014). Quasi-likelihood inference for negative binomial time series models. *Journal of Time Series Analysis*, **35**, 55-78.
- [24] Christou V et Fokianos K (2015a). Estimation and Testing Linearity for Non-Linear Mixed Poisson Autoregressions. *Electronic Journal of Statistics*, **9**, 1357-1377.
- [25] Christou V et Fokianos K (2015b). On Count Time Series Prediction. *Journal of Statistical Computation and Simulation*, **85(2)**, 357-373.
- [26] Cox, D. (1981). Statistical analysis of time-series : some recent developments. *Scand. J. Statist.*, **8**, 93-115.
- [27] Cox, D.R. (1983) . Some remarks on overdispersion, *Biometrika*, **70**, 269-274.
- [28] Davis, R., Dunsmuir, W., and Streett, S. (2003). Observation-driven models for Poisson counts. *Biometrika*, **90**, 777-790.
- [29] Davis, R.A. et Liu, H. (2016). Theory and inference for a class of observation-driven models with application to time series of counts. *Statistica Sinica*, **26**, 1673-1707.
- [30] Davis, R. et Wu, R. (2009). A negative binomial model for time series of counts. *Biometrika*, **96**, 735-749.
- [31] Davis, R.A., Dunsmuir, W.T.M. et Wang, Y. (1999). *Modelling time series of count data*. In Asymptotics, Nonparametrics and Time Series (edt. Subir Ghosh), Marcel Dekker.
- [32] Davis, R.A., Holan, S.H., Lund, R. et Ravishanker, N. (2016). *Handbook of discrete-valued time series*. Chapman and Hall.
- [33] Da Silva, I. M. M. (2005) . *Contributions to the analysis of discrete-valued time series*. PhD thesis, University of Porto, Portugal.
- [34] Dion, J. P., Gauthier, G. et Latour, A. (1995) . Branching processes with immigration and integer-valued time series. *Serdica Math. J.*, **21(2)** , 123-136.
- [35] Diop, M.L. et Kengne, W. (2017). Testing for parameter change in general integer-valued time series. *Journal of Time Series Analysis*, **38**, 880-894.

- [36] Douc, R., Doukhan, P. et Moulines, E. (2013). Ergodicity of observation-driven time series models and consistency of the maximum likelihood estimator. *Stochastic Processes and their Applications*, **123**, 2620-2647.
- [37] Doukhan, P., Fokianos, K., et Tjøstheim, D. (2012). On weak dependence conditions for Poisson autoregressions. *Statistics and Probability Letters*, **82**, 942-948.
- [38] Doukhan, P. et Kengne, W. (2015). Inference and testing for structural change in general Poisson autoregressive models. *Electronic Journal of Statistics*, **9**, 1267-1314.
- [39] Doukhan, P. et Wintenberger, O. (2008). Weakly dependent chains with infinite memory. *Stochastic Processes and their Applications*, **118**, 1997-2013.
- [40] Du Jin-Guan et Li Yuan. (1991). The integer-valued autoregressive ($INAR(p)$) model. *Journal of Time Series Analysis*, **12**, 129-142.
- [41] Dunsmuir W. (2015). Generalized Linear Autoregressive Moving Average Models.” In R Davis, S Holan, R Lund, N Ravishanker (eds.), Handbook of Discrete-Valued Time Series. CRC Monographs.
- [42] Ferland, R., Latour, A. et Oraichi, D. (2006). Integer-valued $GARCH$ process. *Journal of Time Series Analysis*, **27**, 923-942.
- [43] Fokianos, K. (2012). *Count time series models. Handbook in Statistics. Time Series Analysis : Methods and Applications*, **30**, 315-348.
- [44] Fokianos, K. et Kedem, B. (2004). Partial likelihood inference for time series following generalized linear models. *Journal of Time Series Analysis*, **25**, 173-197.
- [45] Fokianos, K., Rahbek, A. et Tjøstheim, D. (2009). Poisson autoregression. *Journal of the American Statistical Association*, **140**, 1430-1439.
- [46] Fokianos, K. et Tjøstheim, D. (2011). Log-linear Poisson autoregression. *Journal of Multivariate Analysis*, **102**, 563-578.
- [47] Fokianos, K, Tjøstheim D. (2012). Nonlinear Poisson autoregression. *Annals of the Institute of Statistical Mathematics*, **64**, 1205-1225.
- [48] Fokianos, K., Neumann, M. H., et al. (2013). A goodness-of-fit test for Poisson count processes. *Electronic Journal of Statistics*, **7**, 793-819.

- [49] Franke J. (2010). Weak dependence of functional INGARCH processes. Technical Report, University of Kaiserslautern.
- [50] Freeland, R. K. (1998). *Statistical analysis of discrete time series with applications to the analysis of workers compensation claims data*. PhD thesis, University of British Columbia, Canada.
- [51] Freeland, R. K. et McCabe. (2004). Analysis of low count time series data by Poisson autoregression. *Journal of Time Series Analysis*, **25** (5), 701-722.
- [52] Freeland, R. K. et McCabe. (2005). Asymptotique properties of CLS estimators in the Poisson $AR(1)$ model. *Statist. Probab. Lett*, **73** (2), 147-153.
- [53] Jung, R. C., Kukuk, M. et Liesenfeld. (2006). Time series of count data : modeling, estimation and diagnostics. *Computational Statistics & Data Analysis*, **51**, 2350-2364.
- [54] Heinen, A. (2003). Modelling time series count data : an autoregressive conditional Poisson model. *Available at SSRN 1117187*.
- [55] Gauthier, G. et Latour, A. (1994). Convergence forte des estimateurs des paramètres d'un processus $GENAR(P)$. *Ann. Sci. Math. Québec*, **18**(1), 49-71.
- [56] Gonçalves, E., Mendes-Lopes, N. et Silva, F. (2015). Infinitely divisible distributions in integer-valued $GARCH$ models. *Journal of Time Series Analysis*, **36**, 503-527.
- [57] Gouriéroux, C., Monfort, A. et Renault, E. (1987). Consistent M-estimators in a Semi-parametric Model. *Document de travail, ENSAE-INSEE*, n°**8706**.
- [58] Gouriéroux, C., Monfort, A. et Trognon, A. (1983a). Testing Nested or Non Nested Hypotheses. *Journal of Econometrics*, **21**, 83-115.
- [59] Gouriéroux, C., Monfort, A. et Trognon, A. (1983b). Estimation par la méthode du Pseudo-maximum de Vraisemblance. *Cahiers du séminaire d'économétrie de M. Malinvaud*, **25**, CNRS, Paris.
- [60] Gouriéroux, C., Monfort, A. et Trognon, A. (1984a). Pseudo maximum likelihood methods : Theory. *Econometrica*, **52**, 681-700.
- [61] Gouriéroux, C., Monfort, A. et Trognon, A. (1984b). Pseudo maximum likelihood methods : Applications to Poisson models. *Econometrica*, **52**, 681-700.

- [62] Grunwald, G., Hyndman, R., Tedesco, L., et Tweedie, R. (2000). Non-Gaussian conditional linear $AR(1)$ models. *Australian and New Zealand Journal of Statistics*, **42**, 479-495.
- [63] Jacobs, P. et Lewis, P. (1978a). Discrete time series generated by mixtures I : correlational and runs properties. *Journal of the Royal Statistical Society Series*, **B40**, 94-105.
- [64] Jacobs, P. et Lewis, P. (1978b). Discrete time series generated by mixtures II : asymptotic properties. *Journal of the Royal Statistical Society Series*, **B40**, 222-228.
- [65] Jacobs, P. et Lewis, P. (1978c). Stationary discrete autoregressive-moving average time series generated by mixtures. *Journal of Time Series Analysis*, **4**, 19-36.
- [66] Kedem, B. et Fokianos, K. (2002). *Regression models for time series analysis*. Wiley, Hoboken, NJ.
- [67] Kengne, W. (2015). Sequential change-point detection in Poisson autoregressive models. *Journal de la Société Française de Statistique*, **156**, 98-112.
- [68] Laroque, G. et Salanie, B. (1987). Estimation of Multi-market Fix-price Models : an Application of Pseudo-maximum Likelihood Methods. *Document de travail, ENSAE-INSEE*, n° **8705**
- [69] Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive process. *Journal of Time Series Analysis*, **19** (4), 439-455.
- [70] Li, W. K. (1994). Time series models based on generalized linear models : some further results. *Biometrics*, **50**, 506-511.
- [71] Liu, H. (2012). *Some models for time series of counts*. PhD thesis, University of Columbia.
- [72] Mac Donald I. L. et Walter Zucchini. (1997). *Hidden Markov and other models for discrete-valued time series*. Monographs on Statistics and Applied Probability **70**. Chapman and Hall, London.
- [73] McCullagh, P. et Nelder, J.A. (1989). *Generalized Linear Models* (2nd ed). London : Chapman & Hall.
- [74] McKenzie, E. (1985a). Contribution to the discussion of Lawrance and Lewis. *J. R. Statist. Soc. B*, **47**, 187-188.

- [75] McKenzie, E. (1985b). Some simple models for discrete variate time series. *Water Resources Bulletin*, **21**, 645-650.
- [76] McKenzie, E. (1986). Autoregressive moving-average processes with negative-binomial and geometric marginal distributions. *Adv . Appl. Prob*, **18**, 679-705.
- [77] McKenzie, E. (1987). Innovation distributions for gamma and negative binomial autoregressions. *Scand. J. Statist*, **14**, 79-85.
- [78] McKenzie, E. (1988a). The distributional structure of finite moving-average processes. *J. Appl. Prob*, **25**, 313-321.
- [79] McKenzie, E. (1988b). Some ARMA models for dependent sequences of Poisson counts. *Adv . Appl. Prob*, **20**, 822-835.
- [80] McKenzie, E. (2003). *Discrete variate time series*, in : Handbook of statistics, Rao, C., Shanbhag, D. (Eds.), Elsevier Science, Amsterdam, 573-606.
- [81] Nelder, J.A. et Wedderburn, R.W. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society, Series A*, **135**, 370-384.
- [82] Neumann MH. 2011. Absolute regularity and ergodicity of Poisson count processes. *Bernoulli*, **17**, 1268-1284.
- [83] Rydberg, T.H.et Shephard, N. (2000). BIN models for trade-by-trade data. Modelling the number of trades in a fixed interval of time. In World Conference Econometric Society, 2000, Seattle. *Contributed Paper 0740*.
- [84] Ristic, M. M. , Bakouch, H. S. et Nastic, A. S. (2009) . A new geometric first-order integer-valued autoregressive (*NGINAR*(1)) process. *J. Statist. Plann. Inference*, **139**(7) , 2218- 2226.
- [85] Samuel, B. (2009) . *Maximum likelihood estimation for INGARCH (p, q) processes*. Project semester, Ecole Polytechnique Fédérale de Lausanne.
- [86] Schweer, S. (2016). A goodness-of-fit test for integer-valued autoregressive processes. *Journal of Time Series Analysis*, **37**, 77 98.
- [87] Silva, M.E. (2015). Modelling time series of counts : an *INAR* approach. *Textos de Matemática*, **47**, 107-121.

- [88] Steutel, F.W. et Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Annals of Probability*, **5**, 893-899.
- [89] Streett, S. (2000). Some observation driven models for time series of counts. Ph.D. thesis, Colorado State University, Department of Statistics.
- [90] Trognon, A. (1984). Pseudo Asymptotic Tests Based on Linear Exponential Families. *Cahiers du séminaire d'économétrie de M.Malinvaud*, **26**, CNRS, Paris.
- [91] Trognon, A. (1987). Les méthodes du Pseudo-maximum de vraisemblance. *Annales d'économie et de Statistique*, **8**.
- [92] Wedderburn, R.W. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika*, **61**, 439-447.
- [93] Weiß, C. H. (2008). Serial dependence and regression of Poisson INARMA models. *J. Statist. Plann. Inference*, **138** (10), 2975- 2990.
- [94] White, H. (1982). Maximum likelihood of misspecified models. *Econometrica*, **50**, 1-25.
- [95] Wooldridge, J.M. (1997). *Quasi-likelihood methods for count data*. In *Handbook of Applied Econometrics*, Volume 2, ed. M.H. Pesaran and P. Schmidt. Oxford : Blackwell, 352-406.
- [96] Wooldridge, J.M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA, MIT Press.
- [97] Zeger, S.L. (1988). A regression model for time series of counts. *Biometrika*, **75**, 621-629.
- [98] Zeger, S.L. et Qaqish, B. (1988). Markov regression models for time series : a quasi-likelihood approach. *Biometrics*, **44**, 1019-1031.
- [99] Zhu, F. (2011). A negative binomial integer-valued *GARCH* model. *Journal of Time Series Analysis*, **32**, 54-67.
- [100] Zhu, F. (2012a). Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued *GARCH* models. *Journal of Mathematical Analysis and its Applications*, **389**, 58-71.
- [101] Zhu, F. (2012b). Zero-inflated Poisson and negative binomial integer-valued *GARCH* models. *Journal of Statistical Planning and Inference*, **142**, 826-839.

- [102] Zhu F. (2012*c*). Modeling time series of counts with COM-Poisson INGARCH models. *Mathematical and Computer Modelling*, **56**, 191-203.
- [103] Zhu, R. et Joe, H. (2006). Modelling count data time series with Markov processes based on binomial thinning. *Journal of Time Series Analysis*, **27** (5), 725-738.