Ann. Sci. Math. Québec **18** (1994), no. 1, 49–71.

CONVERGENCE FORTE DES ESTIMATEURS DES PARAMÈTRES D'UN PROCESSUS GENAR(p)

GENEVIÈVE GAUTHIER ET ALAIN LATOUR

RÉSUMÉ. Dans cet article nous introduisons une généralisation de l'opérateur de STEUTEL et VAN HARN (1979) afin de définir un processus à valeurs entières. Il s'agit d'un processus ayant une structure de corrélation identique à celle du processus autorégressif standard à valeurs réelles. L'estimation des paramètres « α_i » et de la moyenne du bruit de ce modèle est abordée par la méthode des moindres carrés conditionnels. On y démontre la convergence forte de ces estimateurs et des statistiques couramment utilisées dans l'analyse des séries chronologiques.

ABSTRACT. In this paper we introduce a generalization of the STEUTEL and VAN HARN operator (1979). This new operator is used in the definition of an integer valued stochastic process having a correlation structure identical to the correlation structure of the standard real-valued autoregressive process. The problem of the estimation of the parameters "α_i" and of the mean of the noise of this model is tackled using the conditional least squares method. We establish the almost-sure convergence of these estimators and of the statistics currently used in time series analysis.

1. Introduction. Dans cet article, nous introduisons une famille de processus stochastiques dans le but d'enrichir la classe des modèles permettant l'analyse des séries chronologiques à valeurs entières. Nous présentons des résultats relatifs à l'estimation de certains paramètres de ces modèles par la méthode des moindres carrés conditionnels.

L'opérateur de base intervenant dans la définition de plusieurs modèles de séries chronologiques à valeurs entières a été introduit par STEUTEL et VAN HARN (1979). La définition 1.1 précise la nature de cet opérateur.

Définition 1.1. Soit X, une variable aléatoire à valeurs entières non négatives ; $\alpha \in [0,1]$ et $\{Y_i\}_{i\in\mathbb{N}}$, une suite de variables aléatoires indépendantes, identiquement distribuées selon une loi de Bernoulli de paramètre α et indépendantes de X. Alors l'opérateur $\alpha\bullet$ est défini par

$$\alpha \bullet X = \sum_{i=1}^{X} Y_i. \tag{1.1}$$

Reçu le 12 avril 1993 et, sous forme définitive, le 12 octobre 1993.

Recherche faite dans le cadre d'une subvention du Conseil de recherche en sciences naturelles et en génie du Canada

©Association mathématique du Québec

Bien que la notation suggère une simple multiplication, il ne faut pas agir comme tel. Nous donnons à la section 2 les propriétés importantes de cet opérateur.

AL-OSH et ALZAID (1992; 1991; 1990; 1988; 1987) ont suggéré des modèles se voulant semblables aux processus autorégressifs d'ordre p et aux moyennes-mobiles d'ordre q popularisés par BOX et JENKINS et dans lesquels intervient l'opérateur de STEUTEL et VAN HARN. Cependant, le remplacement de la multiplication usuelle par cet opérateur fait apparaître plusieurs contraintes limitant l'utilisation de résultats déjà connus pour les processus autorégressifs à valeurs réelles. Pour contourner le problème, des contraintes au niveau de la définition du processus autorégressif rendent ces modèles moins intéressants qu'ils ne semblaient l'être de prime abord. En effet, la structure de corrélation de leur modèle INAR(p) est semblable à celle d'un ARMA(p, $p-1$).

Les modèles INMA(q) de type moyenne-mobile qu'ils présentent sont prometteurs. Soulignons cependant que ces auteurs n'abordent aucunement la question de l'estimation des paramètres du modèle. Or, il ne nous semble pas que l'opérateur de STEUTEL et VAN HARN puisse permettre l'écriture d'un processus moyenne-mobile sous une forme simple, fonction des valeurs du passé, ce qui rend difficile l'estimation des paramètres. Finalement, notons que plusieurs résultats sont démontrés dans des contextes où les lois marginales sont connues.

Il est cependant possible de définir un processus autorégressif à valeurs entières utilisant l'opérateur de STEUTEL et VAN HARN sans imposer de contraintes supplémentaires aux contraintes liées à la stationnarité.

Définition 1.2. Soit $\{X_t\}_{t\in\mathbb{Z}}$, une suite de variables aléatoires à valeurs entières non négatives ; $\{\varepsilon_t\}_{t\in\mathbb{Z}}$, une suite de variables aléatoires indépendantes et identiquement distribuées à valeurs entières non négatives ; $p \in \mathbb{N}$ et $\{\alpha_k\}_{k\in\{1,2,\ldots,p\}}$, une suite de constantes telles que $\forall k \in \{1, 2, \ldots, p-1\}$, $\alpha_k \in [0, 1]$ et $\alpha_p \in (0, 1]$. Alors $\{X_t\}_{t\in\mathbb{Z}}$ est un processus ENAR(p) (il s'agit d'un acronyme de l'expression : AutoRégressif à valeurs Entières) si

$$X_t = \sum_{k=1}^{p} \alpha_k \bullet X_{t-k} + \varepsilon_t. \tag{1.2}$$

Chacune des suites $\{Y_i\}_{i\in\mathbb{N}}$ sous-entendues dans $\alpha_k \bullet X_{t-k}$, pour $k = 1, 2, \ldots, p$, sont indépendantes entre elles et indépendantes de ε_t.

Cette approche permet d'obtenir une structure de corrélation de type autorégressif. Le problème de l'estimation des paramètres d'un tel processus fut abordé indépendamment par GAUTHIER (1991) et par DU et LI (1991).

JACOBS et LEWIS (1983, 1978a, 1978b, 1978c) et LEWIS (1980) ont, pour leur part, introduit les modèles qu'ils dénomment “Discrete mixed AutoRegressive Moving Average processus” et dont l'acronyme est DARMA. Les processus DARMA sont des combinaisons linéaires aléatoires de variables aléatoires indépendantes et identiquement distribuées. Ces modèles permettent, entre autres choses, de traiter les processus prenant un nombre fini de valeurs et, plus particulièrement, les séries binaires.

D'importantes contributions à l'analyse des séries chronologiques à valeurs discrètes ont été faites par MCKENZIE (1988a, 1988b, 1986, 1985, 1981). Il s'est intéressé aux structures des distributions (autres que celles du premier et des seconds moments) lorsque le processus n'est pas gaussien. Il a aussi traité plusieurs cas particuliers de modèles de type ARMA ayant des lois marginales non gaussiennes.

Le présent article comporte trois volets. À la section 2, nous suggérons une définition plus générale de l'opérateur de STEUTEL et VAN HARN où les variables impliquées dans la somme définissant cet opérateur sont des variables à valeurs entières ayant leurs deux premiers moments finis. Nous n'exigerons pas qu'elles soient de type Bernoulli, ce qui distingue nettement les résultats démontrés ici de ceux de DU et LI (1991). Les propriétés élémentaires de cet opérateur sont établies dans cette section. À la section 3, utilisant ce nouvel opérateur, nous redéfinissons un processus ayant une structure de corrélation identique à celle d'un processus autorégressif standard à valeurs réelles. Nous appellerons GENAR un tel processus. Nous en donnons les deux premiers moments et nous indiquons que la stationnarité d'un tel processus implique des contraintes relatives à certains paramètres du modèle. Ces contraintes sont passées inaperçues dans l'article de DU et LI (1991), bien qu'elles doivent être imposées dans leur contexte, même si le modèle qu'ils proposent est moins général. À la section 4, nous montrons que la moyenne échantillonnale, les autocovariances échantillonnales, de même que les autocorrélations échantillonnales sont des estimateurs fortement convergents des quantités qu'ils estiment. Des résultats semblables apparaissent dans DU et LI (1991) pour le modèle utilisant l'opérateur original de STEUTEL et VAN HARN. De la convergence presque sûre de ces statistiques, nous déduisons finalement la convergence presque sûre des estimateurs des paramètres α_i et de la moyenne de la «composante bruit» du processus GENAR.

2. Généralisation de l'opérateur $\alpha \bullet$.

Définition 2.1. Soit X, une variable aléatoire à valeurs entières non négatives ; Y, une variable aléatoire à valeurs entières non négatives de moyenne finie α et de variance finie λ et $\{Y_i\}_{i \in \mathbb{N}}$, une suite de variables aléatoires indépendantes entre elles, indépendantes de X et distribuées selon la même loi que Y. Alors l'opérateur $\alpha \circ$ est défini par

$$\alpha \circ X = \sum_{i=1}^{X} Y_i. \tag{2.1}$$

Remarque. Nous utilisons la notation $\alpha \circ$ au lieu de $\alpha \bullet$ afin de distinguer l'opérateur généralisé de l'opérateur original.

2.1. Fonction génératrice des probabilités. Dans ce qui suit $\Phi_Z(\cdot)$ dénote la fonction génératrice des probabilités de la variable Z. La proposition 2.1 établit le lien entre la distribution de la variable $\alpha \circ X$ et les distributions de X et de Y. La proposition 2.2 donne la distribution de la variable $\alpha \circ X + \beta \circ Y$, qui est la somme de deux variables indépendantes, transformées indépendamment, c'est-à-dire, que les suites de variables aléatoires intervenant dans $\alpha \circ X$ et dans $\beta \circ Y$ sont indépendantes entre elles. Le lecteur peut facilement démontrer ces deux propositions.

Proposition 2.1. *Soit X et Y telles qu'introduites à la définition 2.1. Alors*

$$\Phi_{\alpha \circ X}(s) = \Phi_X\left(\Phi_Y(s)\right).$$

Proposition 2.2. *Soit X, une variable aléatoire à valeurs entières non négatives ; Y, une variable aléatoire à valeurs entières non négatives de moyenne finie α et de variance finie et W, une variable aléatoire à valeurs entières non négatives de moyenne finie β et de variance finie. Alors*

$$\Phi_{\alpha\circ X+\beta\circ Y}(s) = \Phi_X\left(\Phi_Y(s)\Phi_W(s)\right).$$

Il est facile de voir que si X et Z sont deux variables aléatoires indépendantes à valeurs entières non négatives transformées indépendamment, alors

$$\alpha\circ(X+Z) \stackrel{\mathrm{D}}{=} \alpha\circ X+\alpha\circ Z.$$

La prudence s'impose dans l'utilisation de cet opérateur. Règle générale, nous avons que

i) $\beta\circ[\alpha\circ X] \stackrel{\mathrm{D}}{\neq} [\alpha\beta]\circ X$
ii) $\alpha\circ X+\beta\circ X \stackrel{\mathrm{D}}{\neq} (\alpha+\beta)\circ X$
iii) $\alpha\circ(X+Z) \stackrel{\mathrm{D}}{\neq} \alpha\circ X+\alpha\circ Z$.

2.2. Premier et deuxièmes moments. Soit X et Y satisfaisant la définition 2.1. Puisque $\mathbb{E}(Y)=\alpha<\infty$, et $\mathrm{Var}(Y)=\lambda<\infty$, on peut montrer que

i) $\mathbb{E}(\alpha\circ X)=\alpha\mathbb{E}(X)$
ii) $\mathbb{E}\left((\alpha\circ X)^2\right)=\lambda\mathbb{E}(X)+\alpha^2\mathbb{E}\left(X^2\right)$
iii) $\mathrm{Var}(\alpha\circ X)=\lambda\mathbb{E}(X)+\alpha^2\mathrm{Var}(X)$.

Si de plus Z est une variable aléatoire à valeurs entières non négatives et que les suites de variables aléatoires sous-entendues dans $\beta\circ Z$ et dans $\alpha\circ X$ sont indépendantes, alors

iv) $\mathrm{Cov}(\alpha\circ X,\beta\circ Z)=\alpha\beta\mathrm{Cov}(X,Z)$.

Les résultats démontrés dans BRILLINGER (1969) permettent de retrouver ces résultats et fournissent une méthode de calcul directe basée sur le calcul des cumulants via le calcul des cumulants conditionnels.

3. Généralisation du processus autorégressif. Il est possible de redéfinir le processus autorégressif ENAR(p) avec l'opérateur introduit à la section précédente.

Définition 3.1. Soit $\{X_t\}_{t\in\mathbb{Z}}$, une suite de variables aléatoires à valeurs entières non négatives ; $\{\varepsilon_t\}_{t\in\mathbb{Z}}$, une suite de variables aléatoires indépendantes et identiquement distribuées à valeurs entières non négatives de moyenne finie μ_ε et de variance finie σ_ε^2 ; $p\in\mathbb{N}$ et $\{\alpha_k\}_{k\in\{1,2,\ldots,p\}}$, une suite de constantes telle que pour tout $k\in\{1,2,\ldots,p-1\}$, $\alpha_k\in[0,1)$, $\alpha_p\in(0,1)$ et $\sum_{k=1}^p\alpha_k<1$. Alors $\{X_t\}_{t\in\mathbb{Z}}$ est GENAR(p) si

$$X_t=\sum_{k=1}^p\alpha_k\circ X_{t-k}+\varepsilon_t. \tag{3.1}$$

Toutes les suites $\{Y_i^{(k)}\}_{i\in\mathbb{N}}$ sous-entendues dans $\alpha_k\circ X_{t-k}$, pour $k=1,2,\ldots,p$, sont indépendantes entre elles et indépendantes de ε_t. Elles sont de moyenne finie α_k et de variance finie λ_k. (GENAR est un acronyme de l'expression : AutoRégressif à valeurs ENtières Généralisé.)

Proposition 3.1. *Soit* $\{X_t\}_{t\in\mathbb{Z}}$*, un processus GENAR(p). Alors*

$$\mathbb{E}\left(X_t\right) = \mu_\varepsilon \left(1 - \sum_{k=1}^{p} \alpha_k\right)^{-1}.$$

La démonstration découle directement des résultats de la section 2.2. Puisque $\mathbb{E}\left(X_t\right)$ et μ_ε sont non négatives, nous remarquons que la condition $\sum_{k=1}^{p}\alpha_k < 1$ est nécessaire pour que le processus soit stationnaire au sens large.

Dénotons respectivement par $\gamma(k)$ et $\rho(k)$ l'autocovariance et l'autocorrélation de délai k du processus. Les propositions 3.2 et 3.3 précisent les liens existant entre ces différentes quantités et les paramètres du modèle.

Proposition 3.2. *Soit* $\{X_t\}_{t\in\mathbb{Z}}$*, un processus GENAR(p) stationnaire au sens large. Alors la variance de* X_t *est donnée par*

$$\gamma(0) = \mu_X \sum_{i=1}^{p} \lambda_i + \sum_{i=1}^{p} \alpha_i \gamma(i) + \sigma_\varepsilon^2.$$

La démonstration se fait par simples manipulations algébriques.

Proposition 3.3. *Soit* $\{X_t\}_{t\in\mathbb{Z}}$*, un processus GENAR(p) stationnaire au sens large. Soit* $k \in \mathbb{N}$*,* $k \neq 0$*. Alors l'autocovariance de délai* k *satisfait*

$$\gamma(k) = \sum_{i=1}^{p} \alpha_i \gamma(k-i)$$

et l'autocorrélation de délai k *satisfait*

$$\rho(k) = \begin{cases} 1 & \textit{si } k = 0 \\ \displaystyle\sum_{i=1}^{p} \alpha_i \rho(k-i) & \textit{si } k \geq 1. \end{cases}$$

La démonstration de la proposition 3.3 découle directement des résultats de la section 2.2.

Les autocorrélations satisfont les équations Yule-Walker. Cette structure de corrélation est identique à celle du modèle AR(p) classique à valeurs réelles.

Remarque. En général, il ne semble pas possible d'écrire X_t sous forme d'une somme infinie de termes de la forme $\varphi_j \circ \varepsilon_{t-j}$, à cause de la non-distributivité de l'opérateur $\alpha\circ$ sur l'addition.

Définition 3.2. Soit $\{X_t\}_{t\in\mathbb{Z}}$, un processus stationnaire au sens large et $k \in \mathbb{N}$. Alors, le coefficient de corrélation partielle de délai k, $\phi(k)$, est la k^{e} composante de la solution de $P_k \phi_k = \rho_k$ où

$$P_k = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(k-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(k-2) \\ \rho(2) & \rho(1) & 1 & \cdots & \rho(k-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho(k-1) & \rho(k-2) & \rho(k-3) & \cdots & 1 \end{pmatrix},$$

$$\phi_k = \begin{pmatrix} \phi(1,k) \\ \phi(2,k) \\ \phi(3,k) \\ \vdots \\ \phi(k,k) \end{pmatrix} \quad \text{et} \quad \rho_k = \begin{pmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \\ \vdots \\ \rho(k) \end{pmatrix}.$$

Notons que $\phi(k) = \phi(k,k)$.

Le coefficient de corrélation partielle de délai k d'un processus stationnaire représente le k^{e} coefficient d'un processus de type GENAR(k) dans le cas où nous faisons une approximation du processus stationnaire original par un processus autorégressif d'ordre k. Une conséquence de la proposition 3.3 est que si $\{X_t\}_{t\in\mathbb{Z}}$ est un processus GENAR(p), alors $\phi(k) = 0$, $k > p$.

4. Convergence forte de la moyenne, des autocovariances et des autocorrélations échantillonnales. Soit $X_1, X_2, \dots, X_n$, les valeurs observées d'un processus GENAR(p) stationnaire. Puisque la moyenne, les autocovariances et les autocorrélations de ce type de processus sont reliées entre elles par des relations identiques à celles d'un modèle AR(p) standard, nous utiliserons les mêmes estimateurs que ceux utilisés dans ce dernier cas.

Pour plusieurs résultats de cette section, nous exigerons que

$$\sum_{k=0}^{\infty} |\gamma(k)| < \infty.$$

Pour certains résultats concernant les structures de corrélation, on demandera aussi que

$$\sum_{h=0}^{\infty} |\varphi_k(h)| < \infty$$

où $\varphi_k(h) = \operatorname{Cov}(X_t X_{t+k}, X_{t+h} X_{t+k+h})$. La première hypothèse est très courante dans le domaine des séries chronologiques. Si elle est satisfaite, cela implique que l'autocovariance de délai k tend vers zéro au fur et à mesure que k croît ; ce qui signifie que les observations X_t et X_{t+k} ne sont à peu près pas corrélées lorsqu'elles sont suffisamment éloignées. La deuxième hypothèse demande que les produits $X_t X_{t+k}$ et $X_{t+h} X_{t+k+h}$ ne soient à peu près pas corrélés lorsqu'ils sont suffisamment distants l'un de l'autre.

4.1. Estimation de la moyenne du processus. Comme estimateur de la moyenne du processus, il est naturel de considérer la moyenne arithmétique

$$\overline{X} = \frac{1}{n} \sum_{t=1}^{n} X_t.$$

Évidemment, $\overline{X}$ est sans biais. Le théorème 4.1 en établit la convergence presque sûre vers μ_X.

Théorème 4.1. *Si $\sum_{k=0}^{\infty} |\gamma(k)| < \infty$, alors $\overline{X}$ est un estimateur fortement convergent pour μ_X.*

Ce résultat se démontre en adaptant une preuve de la convergence presque sûre de $\overline{X}$ dans le cas où les observations sont indépendantes donnée par GRIMMETT et STIRZAKER (1983). Le détail est fait à la section 8.

4.2. Estimation des autocovariances. L'estimateur suggéré pour $\gamma(k)$ est

$$\hat{\gamma}(k) = \frac{1}{n} \sum_{t=1}^{n-k} \left(X_t - \overline{X}\right) \left(X_{t+k} - \overline{X}\right)$$

pour $0 \leq k \leq n-1$. Cet estimateur est l'autocovariance échantillonnale de délai k. La proposition 4.2 montre que $\hat{\gamma}(k)$ est un estimateur biaisé de $\gamma(k)$ et que son biais croît en général avec k. Le détail des démonstrations des propositions 4.1 et 4.2, de même que celui du théorème 4.2 est donné à la section 8.

Proposition 4.1. *Pour M fixé et $0 \leq k \leq M$, $\hat{\gamma}(k)$ est asymptotiquement sans biais pour $\gamma(k)$ si $\sum_{k=0}^{\infty} |\gamma(k)| < \infty$.*

Proposition 4.2. *Si $\sum_{k=0}^{\infty} |\gamma(k)| < \infty$, alors, pour un n fixé, le biais de $\hat{\gamma}(k)$ est de l'ordre de n^{-1} et croît en général avec k.*

Théorème 4.2. *Si $\sum_{k=0}^{\infty} |\gamma(k)| < \infty$ et $\sum_{h=0}^{\infty} |\varphi(h)| < \infty$, alors $\hat{\gamma}(k)$ est un estimateur fortement convergent pour $\gamma(k)$.*

4.3. Estimation des autocorrélations. L'autocorrélation échantillonnale de délai k est définie par

$$\hat{\rho}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)}, \text{ pour } 0 \leq k \leq n-1.$$

Puisque l'autocovariance échantillonnale est un estimateur biaisé, il serait vain d'espérer que l'autocorrélation échantillonnale soit sans biais. Cependant, cet estimateur est asymptotiquement sans biais puisque, comme nous l'indique le théorème 4.3, il est fortement convergent pour $\rho(k)$. Le lemme 4.1 est nécessaire à la démonstration du théorème 4.3.

Lemme 4.1. (SERFLING, 1980). *Soit $X_1, X_2, \ldots$ et X des vecteurs aléatoires à valeurs dans $\mathbb{R}^k$ définis sur un espace de probabilité et soit g une fonction définie sur $\mathbb{R}^k$ et mesurable au sens de Borel. Supposons que g soit continue presque sûrement relativement à P_X (l'ensemble de ses discontinuités est de mesure nulle par rapport à la mesure P_X). Alors,*

$$X_n \xrightarrow{\text{p.s.}} X \Longrightarrow g\left(X_n\right) \xrightarrow{\text{p.s.}} g(X).$$

Démonstration. Il s'agit de la partie 1 du théorème de la section 1.7 de SERFLING (1980, p. 24). □

Théorème 4.3. *Si $\gamma(0) > 0$, alors $\hat{\rho}(k)$ est un estimateur fortement convergent pour $\rho(k)$.*

Démonstration. Appliquons le lemme 4.1 avec $g(x, y) = y/x$ qui évaluée au point $\big(\hat{\gamma}(0), \hat{\gamma}(k)\big)$ donne $\hat{\rho}(k)$. La mesurabilité de cette fonction au sens de Borel ne cause aucun problème. Dénotons par E, l'ensemble $\{(x, y) : x = 0\}$. Sur $\mathbb{R}^2 \cap E^c$, g est continue. Puisque $\hat{\gamma}(0)$ est un estimateur fortement convergent pour $\gamma(0) > 0$, $P_X(E) = 0$. Ainsi, g est continue, sauf sur un ensemble de mesure nulle relativement à P_X. $\square$

5. Convergence presque sûre des paramètres α_i d'un processus GENAR(p). Cette section a pour principal but d'établir la convergence forte des estimateurs des moindres carrés des paramètres α_i du modèle GENAR(p). La démonstration est basée sur les résultats de KLIMKO et NELSON (1978). La convergence forte des estimateurs des coefficients d'autocorrélation partielle découle directement de la convergence forte des $\hat{\rho}(k)$.

Soit $\{X_t\}_{t\geq 1}$, un processus stochastique défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P}_\theta)$ dont la distribution dépend d'un vecteur de paramètres inconnus $\theta = \big(\theta_1, \ldots, \theta_p\big)^T$. Ce vecteur varie sur Θ, un ouvert de l'espace euclidien de dimension p.

Nous noterons par $\mathbb{E}_\theta(\cdot)$ et $\mathbb{E}_\theta(\cdot|\cdot)$ l'espérance et l'espérance conditionnelle sous $\mathbb{P}_\theta$. La "vraie" valeur du vecteur de paramètres θ sera notée $\theta^\circ = \big(\theta_1^\circ, \ldots, \theta_p^\circ\big)^T$.

Soit $\{\mathcal{F}_t\}_{t\geq 0}$, une suite de sous-tribus de $\mathcal{F}$ telle que $\mathcal{F}_0 = \{\varnothing, \Omega\}$ et $\mathcal{F}_t$ est engendrée par un sous-ensemble arbitraire de $\{X_1, \ldots, X_t\}$ $\forall t \geq 1$. (Il peut arriver que $\mathcal{F}_t$ soit égale à $\sigma(X_1, \ldots, X_t)$, la plus petite tribu pour laquelle $X_1, \ldots, X_t$ sont mesurables, mais ce n'est pas toujours le cas.) On suppose $\mathbb{E}_\theta\big(|X_t|\big) < \infty$ et ce, pour tout t et pour tout θ. Posons $g(\theta, X_1, \ldots, X_t) = \mathbb{E}_\theta\big(X_{t+1}|\mathcal{F}_t\big)$.

On estime θ en minimisant la somme de carrés conditionnels suivante :

$$Q_n(\theta) = \sum_{t=1}^{n} \big(X_t - g(\theta, X_1, \ldots, X_t)\big)^2$$

par rapport à θ. L'estimateur $\hat{\theta}_n$ de θ est la solution du système à p équations suivant :

$$\forall j \in \{1, 2, \ldots, p\}, \quad \frac{\partial}{\partial \theta_j} Q_n(\theta) = 0.$$

5.1. Estimation des paramètres d'un processus stochastique. La démonstration de la convergence forte de $\hat{\theta}_n$, fait intervenir le développement en série de Taylor de $Q_n(\theta)$ autour de θ°. Il faut supposer que $g(\theta, X_1, \ldots, X_t)$ est continûment différentiable deux fois presque partout dans le voisinage $\|\theta - \theta^\circ\| < \delta^*$ de θ°. Cette dernière condition permet d'utiliser le développement en série de Taylor et assure que la dérivée évaluée en un minimum local de la fonction est nulle.

Posons

$$\nabla Q_n(\theta) = \left[\frac{\partial}{\partial \theta_1} Q_n(\theta), \ldots, \frac{\partial}{\partial \theta_p} Q_n(\theta)\right]^T \text{ et } H_n(\theta) = \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} Q_n(\theta)\right]_{i,j \in \{1,2,\ldots,p\}}.$$

Alors le développement en série de Taylor de $Q_n(\theta)$ autour de θ° est :

$$\begin{aligned} Q_n(\theta) &= Q_n\left(\theta^\circ\right) + \left(\theta - \theta^\circ\right)^T \nabla Q_n\left(\theta^\circ\right) + \frac{1}{2}\left(\theta - \theta^\circ\right)^T H_n\left(\theta^*\right)\left(\theta - \theta^\circ\right) \\ &= Q_n\left(\theta^\circ\right) + \left(\theta - \theta^\circ\right)^T \nabla Q_n\left(\theta^\circ\right) + \frac{1}{2}\left(\theta - \theta^\circ\right)^T H_n\left(\theta^\circ\right)\left(\theta - \theta^\circ\right) \\ &\quad + \frac{1}{2}\left(\theta - \theta^\circ\right)^T \left[H_n\left(\theta^*\right) - H_n\left(\theta^\circ\right)\right]\left(\theta - \theta^\circ\right) \end{aligned}$$

pour un certain point intermédiaire approprié θ^* tel que $\|\theta^* - \theta^\circ\| < \delta^*$.

Remarque. $Q_n(\theta) = \sum_{t=1}^n \left(X_t - g\left(\theta, X_1, \dots, X_{t-1}\right)\right)^2$ est une variable aléatoire. En fait, on devrait noter $Q_n(\omega, \theta)$.

Théorème 5.1. (KLIMKO et NELSON, 1978). *Supposons que*

1) $\forall i, j \in \{1, 2, \dots, p\}$, $\lim_{n\to\infty} \sup_{\delta\to\infty} \frac{1}{n\delta} \left|\left[H_n\left(\omega, \theta^*\right) - H_n\left(\omega, \theta^\circ\right)\right]_{i,j}\right| < \infty$, $\mathbb{P}_{\theta^\circ}$*-p.s.*

2 $\frac{1}{2n} H_n\left(\omega, \theta^\circ\right) \stackrel{n\to\infty}{\longrightarrow} H$, $\mathbb{P}_{\theta^\circ}$*-p.s. où* H *est une matrice de constantes définie positive.*

3) $\frac{1}{n} \nabla Q_n\left(\omega, \theta^\circ\right) \stackrel{n\to\infty}{\longrightarrow} (0, \dots, 0)^T$, $\mathbb{P}_{\theta^\circ}$*-p.s.*

Soit $\varepsilon > 0$ *et* $0 < \delta < \delta^*$. *Alors il existe un ensemble* $N \in \mathcal{F}$, *avec* $\mathbb{P}_{\theta^\circ}(N) > 1 - \varepsilon$ *et un entier* n_0 *tels que, sur* N, $\forall n > n_0$, $\nabla Q_n\left(\omega, \theta^\circ\right) = 0$ *a une solution en* $\hat{\theta}_n(\omega)$ *dans* $B\left(\theta^\circ, \delta\right)$ *(une boule ouverte de rayon* δ *centrée en* θ°*) pour laquelle* $Q_n\left(\omega, \hat{\theta}_n(\omega)\right)$ *est un minimum local.*

Corollaire 5.1. (KLIMKO et NELSON, 1978). *Sous les conditions du théorème 5.1, il existe une suite d'estimateurs* $\{\hat{\theta}_n\}$ *telle que* $\hat{\theta}_n \to \theta^\circ$, $\mathbb{P}_{\theta^\circ}$*-p.s. De plus,* $\forall \varepsilon > 0$, *il existe un ensemble* $N \in \mathcal{F}$ *avec* $\mathbb{P}_{\theta^\circ}(N) > 1 - \varepsilon$ *tel que, sur* N, *pour tout* n *plus grand qu'un certain entier* n_0, $\hat{\theta}_n$ *satisfait* $\nabla Q_n(\omega, \theta) = 0$ *d'où* $Q_n\left(\omega, \hat{\theta}_n\right)$ *est un minimum local.*

5.2. Estimation des paramètres α_i d'un processus autorégressif. Pour démontrer la convergence forte des estimateurs des paramètres α_i d'un processus autorégressif d'ordre p, il suffit de vérifier les trois hypothèses du théorème 5.1.

Nous prendrons comme suite de sous-tribus $\mathcal{F}_0 = \{\varnothing, \Omega\}$, $\mathcal{F}_t = \sigma(X_1, X_2, \dots, X_p)$, pour $1 \le t \le p$ et $\mathcal{F}_t = \sigma\left(X_1, X_2, \dots, X_t\right)$, pour $t > p$. Seuls les p premiers termes de cette suite ne répondent pas aux exigences posées sur la suite de sous-tribus construite au début de la section 5, donc les résultats énoncés à la section 5.1 peuvent quand même s'appliquer.

Pour le cas particulier qui nous intéresse, nous avons

$$g\left(\alpha, \mu_\varepsilon, X_1, \dots, X_t\right) = \sum_{i=1}^p \alpha_i X_{t+1-i} + \mu_\varepsilon, \quad t \ge p,$$

car $\mathbb{E}(\alpha \circ X \mid X) = \alpha X$ puisque, conditionnellement à X, $\alpha \circ X$ est une somme de X variables aléatoires indépendantes de moyenne α. Par conséquent, pour $n > p$,

$$Q_n\left(\alpha, \mu_\varepsilon\right) = \sum_{t=p+1}^n \left(X_t - \sum_{i=1}^p \alpha_i X_{t-i} - \mu_\varepsilon\right)^2. \tag{5.1}$$

Ainsi,

$$\nabla Q_n(\alpha,\mu_\varepsilon) = \begin{pmatrix} \dfrac{\partial}{\partial \alpha_1} Q_n(\alpha,\mu_\varepsilon) \\ \vdots \\ \dfrac{\partial}{\partial \alpha_p} Q_n(\alpha,\mu_\varepsilon) \\ \dfrac{\partial}{\partial \mu_\varepsilon} Q_n(\alpha,\mu_\varepsilon) \end{pmatrix} = \begin{pmatrix} -2\displaystyle\sum_{t=p+1}^{n} X_{t-1}\left(X_t - \sum_{i=1}^{p}\alpha_i X_{t-i} - \mu_\varepsilon\right) \\ \vdots \\ -2\displaystyle\sum_{t=p+1}^{n} X_{t-p}\left(X_t - \sum_{i=1}^{p}\alpha_i X_{t-i} - \mu_\varepsilon\right) \\ -2\displaystyle\sum_{t=p+1}^{n} \left(X_t - \sum_{i=1}^{p}\alpha_i X_{t-i} - \mu_\varepsilon\right) \end{pmatrix}$$

et $H_n(\alpha,\mu_\varepsilon)$

$$= \begin{pmatrix} 2\displaystyle\sum_{t=p+1}^{n} X_{t-1}^2 & 2\displaystyle\sum_{t=p+1}^{n} X_{t-1}X_{t-2} & \cdots & 2\displaystyle\sum_{t=p+1}^{n} X_{t-1}X_{t-p} & 2\displaystyle\sum_{t=p+1}^{n} X_{t-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 2\displaystyle\sum_{t=p+1}^{n} X_{t-1}X_{t-p} & 2\displaystyle\sum_{t=p+1}^{n} X_{t-2}X_{t-p} & \cdots & 2\displaystyle\sum_{t=p+1}^{n} X_{t-p}^2 & 2\displaystyle\sum_{t=p+1}^{n} X_{t-p} \\ 2\displaystyle\sum_{t=p+1}^{n} X_{t-1} & 2\displaystyle\sum_{t=p+1}^{n} X_{t-2} & \cdots & 2\displaystyle\sum_{t=p+1}^{n} X_{t-p} & \displaystyle\sum_{t=p+1}^{n} 2 \end{pmatrix}$$

Dans ce qui suit, θ° désigne $\left(\alpha^\circ,\mu_\varepsilon^\circ\right)$ qui est le vecteur des véritables valeurs des paramètres du modèle. Le symbole $\mathbb{P}_{\theta^\circ}$ désigne la loi du processus pour ces valeurs des paramètres et $\mathbb{E}$ représente l'espérance relativement à la mesure $\mathbb{P}_{\theta^\circ}$.

Le lemme 5.1 vérifie la première condition du théorème 5.1.

Lemme 5.1. $\forall i,j \in \{1,2,\ldots,p\}$

$$\lim_{n\to\infty}\sup_{\delta\to 0}\frac{1}{n\delta}\left|\left[H_n\left(\omega,\alpha^*,\mu_\varepsilon^*\right) - H_n\left(\omega,\alpha^\circ,\mu_\varepsilon^\circ\right)\right]_{i,j}\right| < \infty, \quad \mathbb{P}_{\theta^\circ}\text{-}p.s.$$

Démonstration. Il suffit de remarquer que $H_n\left(\omega,\alpha^*,\mu_\varepsilon^*\right) = H_n\left(\omega,\alpha^\circ,\mu_\varepsilon^\circ\right)$. □

La démonstration de la troisième condition du théorème 5.1 fait intervenir le théorème 5.2 relatif aux martingales.

Théorème 5.2. (HALL et HEYDE, 1980). *Soit* $\left\{S_n = \sum_{t=1}^{n} W_t, \mathcal{F}_n\right\}$ *une martingale et* $\{U_n\}_{n\geq 1}$ *une suite non décroissante de variables aléatoires positives telles que* U_n *soit* $\mathcal{F}_{n-1}$*-mesurable pour tout* n*. Alors, si* $1 \leq s \leq 2$*,*

$$\lim_{n\to\infty}\frac{S_n}{U_n} = 0, \quad p.s.$$

sur l'ensemble $\left\{\lim_{n\to\infty} U_n = \infty \text{ et } \displaystyle\sum_{t=1}^{\infty}\frac{1}{U_t^s}\mathbb{E}\left(|W_t|^s \mid \mathcal{F}_{t-1}\right) < \infty\right\}$.

Une démonstration de ce théorème se trouve dans HALL et HEYDE (1980, pp. 35–36).

Les lemmes qui suivent ont pour objectif la vérification des conditions du théorème 5.2 lorsque $S_n = [\nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right)]_i$, pour $1 \leq i \leq p+1$, $U_n = n$ et $s = 2$. Cette application particulière du théorème 5.2 nous permet de montrer que chacune des composantes de $n^{-1}[\nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right)]$ converge presque sûrement vers 0 et, par conséquent, que $n^{-1}\nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right) \stackrel{n\to\infty}{\longrightarrow} (0, \ldots, 0)^T$ $\mathbb{P}_{\theta^\circ}$-p.s. (qui est la troisième condition du théorème 5.1). Le lemme 5.2 établit le fait que nous travaillons avec des martingales ; le lemme 5.3 montre que pour chacune des composantes du vecteur $\nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right)$, l'ensemble dont il est question au théorème 5.2 est de probabilité 1. Le détail de ces démonstrations se retrouve à la section 8.

Lemme 5.2. *Chaque composante du vecteur $\nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right)$ est une $\mathcal{F}_n$-martingale pour $\mathbb{P}_{\theta^\circ}$.*

Lemme 5.3.

$$\mathbb{P}_{\theta^\circ}\left\{\lim_{n\to\infty} n = \infty \text{ et } \sum_{n=1}^{\infty}\frac{1}{n^2}\mathbb{E}\left(\left|X_n - \sum_{j=1}^{p}\alpha_j^\circ X_{n-j} - \mu_\varepsilon^\circ\right|^2 \middle| \mathcal{F}_{n-1}\right) < \infty\right\} = 1.$$

Lemme 5.4. *Pour $1 \leq i \leq p$, si $\forall h$, $\varphi_0(h) < \infty$,*

$$\mathbb{P}_{\theta^\circ}\left\{\lim_{n\to\infty} n = \infty \text{ et } \sum_{n=1}^{\infty}\frac{1}{n^2}\mathbb{E}\left(\left|X_{n-i}\left(X_n - \sum_{j=1}^{p}\alpha_j^\circ X_{n-j} - \mu_\varepsilon^\circ\right)\right|^2 \middle| \mathcal{F}_{n-1}\right) < \infty\right\} = 1.$$

Corollaire 5.2. *Si $\forall h$, $\varphi_0(h) < \infty$, alors*

$$\frac{1}{n}\nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right) \stackrel{n\to\infty}{\longrightarrow} (0, \ldots, 0)^T, \quad \mathbb{P}_{\theta^\circ}\text{-p.s.}$$

Démonstration. Ceci est une conséquence immédiate du théorème 5.2 et des lemmes 5.2, 5.3 et 5.4. □

Rappelons que

$$\Gamma = \begin{pmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \cdots & \gamma(p-1) \\ \gamma(1) & \gamma(0) & \gamma(1) & \cdots & \gamma(p-2) \\ \gamma(2) & \gamma(1) & \gamma(0) & \cdots & \gamma(p-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma(p-1) & \gamma(p-2) & \gamma(p-3) & \cdots & \gamma(0) \end{pmatrix}.$$

Lemme 5.5. *Si $\gamma(0) > 0$ et si $\sum_{h=0}^{\infty}|\varphi_k(h)| < \infty$, alors $(2n)^{-1}H_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right) \stackrel{n\to\infty}{\longrightarrow} H$, $\mathbb{P}_{\theta^\circ}$-p.s., où H est une matrice de constantes définie positive.*

Démonstration. D'une part, $(2n)^{-1}H_n\left(\alpha, \mu_\varepsilon^\circ\right)$ est égal à

$$\begin{pmatrix} \frac{1}{n}\sum_{t=p+1}^{n} X_{t-1}^2 & \frac{1}{n}\sum_{t=p+1}^{n} X_{t-1}X_{t-2} & \cdots & \frac{1}{n}\sum_{t=p+1}^{n} X_{t-1}X_{t-p} & \frac{1}{n}\sum_{t=p+1}^{n} X_{t-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{n}\sum_{t=p+1}^{n} X_{t-1}X_{t-p} & \frac{1}{n}\sum_{t=p+1}^{n} X_{t-2}X_{t-p} & \cdots & \frac{1}{n}\sum_{t=p+1}^{n} X_{t-p}^2 & \frac{1}{n}\sum_{t=p+1}^{n} X_{t-p} \\ \frac{1}{n}\sum_{t=p+1}^{n} X_{t-1} & \frac{1}{n}\sum_{t=p+1}^{n} X_{t-2} & \cdots & \frac{1}{n}\sum_{t=p+1}^{n} X_{t-p} & \sum_{t=p+1}^{n} \frac{1}{n} \end{pmatrix}.$$

D'autre part, dans la démonstration du théorème 4.2, on constate que $n^{-1}\sum_{t=1}^{n-k} X_t X_{t+k}$ converge presque sûrement vers $\gamma(k)+\mu_X^2$. Par conséquent, H est la matrice suivante :

$$\begin{pmatrix} \gamma(0)+\mu_X^2 & \gamma(1)+\mu_X^2 & \gamma(2)+\mu_X^2 & \cdots & \gamma(p-1)+\mu_X^2 & \mu_X \\ \gamma(1)+\mu_X^2 & \gamma(0)+\mu_X^2 & \gamma(1)+\mu_X^2 & \cdots & \gamma(p-2)+\mu_X^2 & \mu_X \\ \gamma(2)+\mu_X^2 & \gamma(1)+\mu_X^2 & \gamma(0)+\mu_X^2 & \cdots & \gamma(p-3)+\mu_X^2 & \mu_X \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma(p-1)+\mu_X^2 & \gamma(p-2)+\mu_X^2 & \gamma(p-3)+\mu_X^2 & \cdots & \gamma(0)+\mu_X^2 & \mu_X \\ \mu_X & \mu_X & \mu_X & \cdots & \mu_X & 1 \end{pmatrix}.$$

Soit

$$G = \left(\begin{array}{c|c} I_p & 0_p \\ \hline -\mu_X 1_p' & 1 \end{array}\right)$$

où 1_p est le vecteur de longueur p dont toutes les composantes sont égales à 1, tandis que 0_p est le vecteur dont toutes les composantes sont nulles. On se convainc assez facilement que G est inversible et

$$G'HG = \left(\begin{array}{c|c} \Gamma & 0_p \\ \hline 0_p' & 1 \end{array}\right)$$

Cette dernière matrice est définie positive, puisque $\gamma(0) > 0$ (BROCKWELL et DAVIS, 1991, p. 167). □

Théorème 5.3. *Si $\gamma(0) > 0$ et si $\sum_{h=0}^{\infty} |\varphi_k(h)| < \infty$, alors, la solution de $\nabla Q_n(\alpha, \mu_\varepsilon) = 0$ génère des estimateurs fortement convergents pour $\alpha_1, \alpha_2, \ldots, \alpha_p$ et μ_ε.*

Démonstration. Les lemmes 5.2 et 5.5 ainsi que le corollaire 5.1 nous assurent qu'il est possible d'utiliser le théorème 5.1 et le corollaire 5.1. □

5.3. Estimations des coefficients d'autocorrélation partielle. Comme le coefficient d'autocorrélation partielle de délai k, $\phi(k)$, est la k^{e} composante de la solution de $P_k\phi_k = \rho_k$ où

$$P_k = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(k-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(k-2) \\ \rho(2) & \rho(1) & 1 & \cdots & \rho(k-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho(k-1) & \rho(k-2) & \rho(k-3) & \cdots & 1 \end{pmatrix},$$

$$\phi_k = \begin{pmatrix} \phi(1,k) \\ \phi(2,k) \\ \phi(3,k) \\ \vdots \\ \phi(k,k) \end{pmatrix} \quad et \quad \rho_k = \begin{pmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \\ \vdots \\ \rho(k) \end{pmatrix},$$

nous suggérons comme estimateur du coefficient de corrélation partielle de délai k la dernière composante de la solution de ce même système en remplaçant les $\rho(j)$ par $\hat{\rho}(j)$. Puisqu'il y a convergence forte des $\hat{\rho}(j)$ vers leur valeur théorique respective et que la solution du système d'équations ci-haut est une fonction des $\hat{\rho}(j)$ mesurable au sens de Borel, nous avons aussi la convergence forte des estimateurs des coefficients d'autocorrélation partielle. Ceci est dû au fait que nous avons supposé que $\gamma(0) > 0$, ce qui entraîne que la matrice des variance-covariance Γ de k observations consécutives est définie positive.

6. Conclusion. Nous avons proposé une extension naturelle au modèle INAR(p). Il serait intéressant d'élargir encore la classe des processus à valeurs discrètes en considérant des processus du genre ARMA de la forme

$$X_t = \sum_{k=1}^{p} \alpha_k \circ X_{t-k} + \varepsilon_t - \sum_{k=1}^{q} \beta_k \circ \varepsilon_{t-k},$$

définis à partir de l'opérateur généralisé de STEUTEL et VAN HARN. L'ensemble de ces résultats pourrait être la base d'une méthodologie comparable à celle de BOX et JENKINS (1976) applicable à des séries chronologiques à valeurs discrètes.

7. Remerciements. Les auteurs aimeraient remercier les arbitres pour leurs conseils et leurs remarques qui ont conduit à cette version finale. Ils remercient aussi le décanat des études avancées et de la recherche de l'UQAM qui a permis la publication, sous forme de rapport interne de recherche, d'une version préliminaire de cet article. Finalement, Alain Latour remercie particulièrement Johanne Corbeil et Jefferey L. Terry pour leur support durant son année sabbatique, période durant laquelle la dernière version de cet article fut rédigée.

8. Appendice.

8.1. Preuve du théorème 4.1. Soit $(\Omega, \mathcal{F}, \mathrm{Pr})$, un espace de probabilité. Pour cette démonstration, nous dénoterons par $\overline{X}_n(\omega)$ la moyenne arithmétique $n^{-1}\sum_{t=1}^{n} X_t = n^{-1}S_n(\omega)$. Considérons dans un premier temps une sous-suite $\{n_i^{-1}S_{n_i}\}$, avec $n_i = i^2$. Par l'inégalité de Tchebycheff, puisque $\overline{X}_n(\omega)$ est sans biais, pour tout $\varepsilon > 0$,

$$\mathrm{Pr}\left(\frac{1}{i^2}\left|S_{i^2}(\omega) - i^2\mu_X\right| > \varepsilon\right) \le \frac{\mathrm{Var}\left(S_{i^2}(\omega)\right)}{i^4\varepsilon^2} = \frac{1}{i^2\varepsilon^2}\sum_{k=-i^2+1}^{i^2-1}\left(1 - \frac{|k|}{i^2}\right)\gamma(k),$$

en utilisant le fait que Var $\left(\overline{X}_n\right) = \frac{1}{n}\sum_{r=-(n-1)}^{n-1}\left(1-\frac{|r|}{n}\right)\gamma(r)$ (voir le théorème 8.2.3 de ANDERSON (1971)), pour établir la dernière égalité. En sommant sur i, on obtient,

$$\sum_{i=1}^{\infty}\Pr\left(\frac{1}{i^2}\left|S_{i^2}(\omega)-i^2\mu_X\right|>\varepsilon\right)\leq\frac{1}{\varepsilon^2}\sum_{i=1}^{\infty}\frac{1}{i^2}\sum_{k=-i^2+1}^{i^2-1}\left(1-\frac{|k|}{i^2}\right)\gamma(k).$$

Quelques manipulations algébriques directes montrent que le second terme peut être borné par

$$\frac{\pi^2}{6\varepsilon^2}\sum_{k=-\infty}^{\infty}|\gamma(k)|<\infty.$$

Par conséquent, $\sum_{i=1}^{\infty}\Pr\left(\frac{1}{i^2}|S_{i^2}(\omega)-i^2\mu_X|>\varepsilon\right)<\infty$. Par le lemme de Borel-Cantelli,

$$\Pr\left(\lim_{i\to\infty}\sup\left\{\omega:\left|\overline{X}_{i^2}(\omega)-\mu_X\right|>\varepsilon\right\}\right)=0.$$

Comme ε est arbitraire, on peut conclure que

$$\frac{1}{i^2}S_{i^2}(\omega)\xrightarrow{\text{p.s.}}\mu_X.$$

Dans un second temps, considérons m tel que $i^2\leq m\leq(i+1)^2$. Puisque la variable X_t est non négative, on a que

$$S_{i^2}(\omega)\leq S_m(\omega)\leq S_{(i+1)^2}(\omega)$$

d'où

$$\frac{i^2}{(i+1)^2}\frac{1}{i^2}S_{i^2}(\omega)\leq\frac{1}{m}S_m(\omega)\leq\frac{(i+1)^2}{i^2}\frac{1}{(i+1)^2}S_{(i+1)^2}(\omega).$$

Ainsi, lorsque m tend vers l'infini,

$$\frac{1}{m}S_m(\omega)\xrightarrow{\text{p.s.}}\mu_X.\quad\square$$

8.2. Preuve des propositions 4.1 et 4.2. Pour la démonstration de la proposition 4.1, nous avons besoin des résultats suivants :

Lemme 8.1. *Soit $\{s_k\}_{k\in\mathbb{N}}$, une suite de nombres réels ou complexes. Si $\sum_{k=1}^{\infty}s_k$ converge, alors*

$$\lim_{\tau\to\infty}\sum_{k=1}^{\tau-1}\left(1-\frac{k}{\tau}\right)s_k=\sum_{k=1}^{\infty}s_k.$$

La démonstration de ce lemme se trouve dans ANDERSON (1971, p. 460).

Corollaire 8.1. *Si* $\sum_{k=1}^{\infty}|\gamma(k)|<\infty$ *alors* $\lim_{n\to\infty}\sum_{k=-n}^{n}\left(1-\frac{|k|}{n}\right)\gamma(k)=\sum_{k=-\infty}^{\infty}\gamma(k).$

Abordons maintenant la preuve des propositions 4.1 et 4.2. Pour $k=0$, en développant

$$\mathbb{E}\left(\hat{\gamma}(0)\right)=\mathbb{E}\left(\frac{1}{n}\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2\right)$$

on trouve

$$\mathbb{E}\left(\hat{\gamma}(0)\right) = \frac{n-1}{n}\gamma(0) - \sum_{i=1}^{n-1} \frac{2(n-i)}{n^2}\gamma(i).$$

Par conséquent,

$$\lim_{n\to\infty} \mathbb{E}\left(\hat{\gamma}(0)\right) = \lim_{n\to\infty} \left[\frac{n-1}{n}\gamma(0) - \frac{2}{n}\sum_{i=1}^{n-1}\left(1-\frac{i}{n}\right)\gamma(i)\right] = \gamma(0)$$

à cause du corollaire 8.1. Pour $k > 0$, on montre dans un premier temps que

$$\begin{aligned}\mathbb{E}\left(\hat{\gamma}(k)\right) = \frac{1}{n}\Bigg[&\sum_{t=1}^{n-k}\gamma(k) - \frac{1}{n}\sum_{t=1}^{n-k}\sum_{i=1}^{n}\gamma(t-i)\\ &-\frac{1}{n}\sum_{t=1}^{n-k}\sum_{i=1}^{n}\gamma(t+k-i) + \frac{1}{n^2}\sum_{t=1}^{n-k}\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma(i-j)\Bigg].\end{aligned}$$

Pour les différents termes de l'expression entre crochets, nous avons :

$$\begin{aligned}
S_0 &:= \sum_{t=1}^{n-k}\gamma(k) = (n-k)\gamma(k)\\
S_1 &:= \sum_{t=1}^{n-k}\sum_{i=1}^{n}\gamma(t-i) = \sum_{i=1}^{n-k-1}(n-k-i)\gamma(i) + (n-k)\sum_{i=0}^{k}\gamma(i) + \sum_{i=k+1}^{n-1}(n-i)\gamma(i)\\
S_2 &:= \sum_{t=1}^{n-k}\sum_{i=1}^{n}\gamma(t+k-i)\\
&= \sum_{i=k+1}^{n-1}(n-i)\gamma(i) + (n-k)\sum_{i=0}^{k}\gamma(i) + \sum_{i=1}^{n-k-1}(n-k-i)\gamma(i)\\
S_3 &:= \sum_{t=1}^{n-k}\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma(i-j) = 2(n-k)\sum_{i=1}^{n-1}(n-i)\gamma(i) + n(n-k)\gamma(0).
\end{aligned}$$

Substituant ces expressions dans celle de $\mathbb{E}\left(\hat{\gamma}(k)\right)$, après quelques manipulations algébriques et en faisant tendre n vers l'infini, nous trouvons,

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\left(\hat{\gamma}(k)\right) = \lim_{n\to\infty}\Bigg[&\frac{n-k}{n}\gamma(k) - \frac{2}{n^2}\Bigg(\sum_{i=1}^{n-k-1}(n-k-i)\gamma(i) + (n-k)\sum_{i=0}^{k}\gamma(i)\\ &+ \sum_{i=k+1}^{n-1}(n-i)\gamma(i)\Bigg) + \frac{2(n-k)}{n^3}\sum_{i=1}^{n-1}(n-i)\gamma(i) + \frac{(n-k)}{n^2}\gamma(0)\Bigg]\end{aligned}$$

$$= \gamma(k) + \lim_{n\to\infty}\left[\frac{-k}{n}\gamma(k) - \frac{(n-k)}{n^2}\gamma(0) - \frac{2}{n}\sum_{i=1}^{k}\gamma(i) + \frac{2k}{n^2}\sum_{i=1}^{k}\gamma(i) + \frac{2k}{n^2}\sum_{i=1}^{n-k-1}\frac{i}{n}\gamma(i) - \frac{2k}{n^2}\sum_{i=k+1}^{n-1}\left(1-\frac{i}{n}\right)\gamma(i)\right].$$

Nous remarquons alors que

$$\lim_{n\to\infty}\left[k\gamma(k) + \frac{(n-k)}{n}\gamma(0) + 2\sum_{i=1}^{k}\gamma(i) - \frac{2k}{n}\sum_{i=1}^{k}\gamma(i) - \frac{2k}{n}\sum_{i=1}^{n-k-1}\frac{i}{n}\gamma(i) + \frac{2k}{n}\sum_{i=k+1}^{n-1}\left(1-\frac{i}{n}\right)\gamma(i)\right] = k\gamma(k) + \sum_{i=-k}^{k}\gamma(i).$$

Des deux dernières égalités, nous concluons que $\hat{\gamma}(k)$ est asymptotiquement sans biais pour $\gamma(k)$ et que le biais, d'une manière générale, croît avec k. De plus, le biais est de l'ordre de $o\left(n^{-1}\right)$.

8.3. Preuve du théorème 4.2. Dans la démonstration du théorème 4.2, nous utilisons les résultats du lemme 8.2.

Lemme 8.2. *$\overline{X}_n \overset{n\to\infty}{\longrightarrow} \mu_X$ p.s. implique que*

1) $\lim_{n\to\infty} \overline{X}_n \frac{1}{n}\sum_{t=1}^{n-k} X_{t+k} = \mu_X^2$ *p.s.*
2) $\lim_{n\to\infty} \overline{X}_n \frac{1}{n}\sum_{t=1}^{n-k} X_{t} = \mu_X^2$ *p.s.*
3) $\lim_{n\to\infty} \frac{n-k}{n}\overline{X}_n^2 = \mu_X^2$ *p.s.*

Démonstration. Nous allons démontrer la partie 1 ; les deux autres se prouvent d'une manière analogue. Notons que

$$\lim_{n\to\infty} \overline{X}_n \frac{1}{n}\sum_{t=1}^{n-k} X_{t+k} = \lim_{n\to\infty} \overline{X}_n^2 - \lim_{n\to\infty} \overline{X}_n \frac{1}{n}\sum_{t=1}^{k} X_t.$$

Traitons les deux termes indépendamment. Premièrement, puisque par le théorème 4.1

$$\overline{X}_n \overset{n\to\infty}{\longrightarrow} \mu_X \text{ p.s.},$$

en appliquant le lemme 4.1 avec $g\left(\overline{X}_n\right) = \overline{X}_n^2$ nous concluons que

$$\overline{X}_n^2 \overset{n\to\infty}{\longrightarrow} \mu_X^2 \text{ p.s.}$$

Deuxièmement, puisque $\sum_{t=1}^{k} X_t$ ne dépend pas de n, $n^{-1}\sum_{t=1}^{k} X_t \overset{n\to\infty}{\longrightarrow} 0$, p.s. Sachant que $\overline{X}_n \overset{\text{p.s.}}{\longrightarrow} \mu_X$, nous concluons que

$$\overline{X}_n n^{-1}\sum_{t=1}^{k} X_t \overset{n\to\infty}{\longrightarrow} 0 \text{ p.s.}$$

en appliquant le lemme 4.1 avec

$$g\left(\overline{X}_n, \frac{1}{n}\sum_{t=1}^{k} X_t\right) = \overline{X}_n \frac{1}{n}\sum_{t=1}^{k} X_t. \quad \square$$

Procédons maintenant à la démonstration du théorème 4.2. En supposant que $n > k$, nous avons

$$\hat{\gamma}(k) = \frac{1}{n}\sum_{t=1}^{n-k} X_t X_{t+k} - \overline{X}_n \frac{1}{n}\sum_{t=1}^{n-k} X_t - \overline{X}_n \frac{1}{n}\sum_{t=1}^{n-k} X_{t+k} + \frac{n-k}{n}\overline{X}_n^2.$$

À cause des lemmes 4.1 et 8.2, démontrer que $\hat{\gamma}(k) \overset{n\to\infty}{\longrightarrow} \gamma(k)$ p.s. se réduit à démontrer que $n^{-1}\sum_{t=1}^{n-k} X_t X_{t+k}$ converge presque sûrement vers $\gamma(k) + \mu_X^2$.

Notons que $n^{-1}\sum_{t=1}^{n-k} X_t X_{t+k}$ est asymptotiquement sans biais pour $\gamma(k) + \mu_X^2$ car

$$\mathbb{E}\left(\frac{1}{n}\sum_{t=1}^{n-k} X_t X_{t+k}\right) = \frac{n-k}{n}\left(\gamma(k) + \mu_X^2\right) \overset{n\to\infty}{\longrightarrow} \gamma(k) + \mu_X^2.$$

Quelques manipulations algébriques nous indiquent que

$$\mathbb{E}\left(\frac{1}{n}\sum_{t=1}^{n-k} X_t X_{t+k}\right)^2 = \frac{1}{n^2}\sum_{t=1}^{n-k}\sum_{u=1}^{n-k} \varphi_k(t-u) + \left(\frac{n-k}{n}\right)^2 \left(\gamma(k) + \mu_X^2\right)^2.$$

Par conséquent,

$$\operatorname{Var}\left(\frac{1}{n}\sum_{t=1}^{n-k} X_t X_{t+k}\right) = \frac{n-k}{n^2}\sum_{h=-(n-k)+1}^{(n-k)-1}\left(1 - \frac{|h|}{n-k}\right)\varphi_k(h).$$

Par le corollaire 8.1,

$$\lim_{n\to\infty} \operatorname{Var}\left(\frac{1}{n}\sum_{t=1}^{n-k} X_t X_{t+k}\right) = \left(\lim_{n\to\infty}\frac{n-k}{n^2}\right)\left(\sum_{h=-\infty}^{\infty}\varphi_k(h)\right) = 0.$$

Soit $T_n = \sum_{t=1}^{n-k} X_t X_{t+k}$. La preuve se termine en utilisant un argument analogue à celui utilisé dans la démonstration du théorème 4.1 afin de démontrer que $n^{-1}T_n$ converge presque sûrement vers $\gamma(k) + \mu_X^2$. $\square$

8.5. Preuve des lemmes 5.2, 5.3 et 5.4.

Preuve du lemme 5.2. Soit $\{\mathcal{F}_n\}_{n\geq 1}$, une filtration telle que $\mathcal{F}_t = \sigma\left(X_1, X_2, \ldots, X_p\right)$ pour $1 \leq t \leq p$ et $\mathcal{F}_t = \sigma\left(X_1, X_2, \ldots, X_t\right)$ pour $t > p$, $\sigma\left(X_1, X_2, \ldots, X_p\right)$ étant la plus petite tribu rendant mesurables $X_1, X_2, \ldots, X_t$. Nous voulons montrer que chaque composante du vecteur $\nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right)$ est une $\mathcal{F}_n$-martingale. En effet, nous avons que

1) $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_n \subset \cdots$.

2) Pour $1 \leq i \leq p+1$, $\nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right)$ est $\mathcal{F}_n$-mesurable.
3) Pour $1 \leq i \leq p+1$, $\nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right)$ est intégrable. En effet, pour $1 \leq i \leq p$, on montre que l'expression

$$\mathbb{E}\left(\left|\left[\nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right)\right]_i\right|\right) = 2\mathbb{E}\left(\left|\sum_{t=p+1}^{n} X_{t-i}\left(X_t - \sum_{j=1}^{p} \alpha_j^\circ X_{t-j} - \mu_\varepsilon^\circ\right)\right|\right)$$

est bornée par

$$4(n-p)\left(\gamma(i) + \mu_X^2\right) < \infty.$$

Pour $i = p+1$, on montre que l'expression

$$\mathbb{E}\left(\left|\left[\nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right)\right]_i\right|\right) = 2\mathbb{E}\left(\left|\sum_{t=p+1}^{n}\left(X_t - \sum_{j=1}^{p} \alpha_j^\circ X_{t-j} - \mu_\varepsilon^\circ\right)\right|\right)$$

est bornée par

$$4(n-p)\mu_X < \infty.$$

4) Quelques manipulations algébriques directes mettent en évidence le fait que

$$\mathbb{E}\left(\nabla Q_{n+1}\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right) \mid \mathcal{F}_n\right) = \nabla Q_n\left(\omega, \alpha^\circ, \mu_\varepsilon^\circ\right)$$

pour $1 \leq i \leq p+1$. $\square$

Preuve du lemme 5.3. D'une part

$$\lim_{n\to\infty} U_n = \lim_{n\to\infty} n = \infty.$$

D'autre part, puisque $\mathbb{E}\left[\left(\alpha_j \circ X_{n-j} - \alpha_j X_{n-j}\right)^2 \mid X_{n-j}\right] = \mathrm{Var}[\alpha_j \circ X_{n-j}] = \lambda_j X_{n-j}$, on voit facilement que

$$\mathbb{E}\left(\left|X_n - \sum_{j=1}^{p} \alpha_j^\circ X_{n-j} - \mu_\varepsilon^\circ\right|^2 \middle| \mathcal{F}_{n-1}\right) = \mathrm{Var}[X_n \mid \mathcal{F}_{n-1}] = \sum_{j=1}^{p} \lambda_j X_{n-j} + \sigma_\varepsilon^2$$
$$\leq \max_{1\leq k\leq p}\{\lambda_k\} \sum_{j=1}^{p} X_{n-j} + \sigma_\varepsilon^2.$$

Par conséquent,

$$\sum_{n=1}^{\infty} \frac{1}{n^2}\mathbb{E}\left(\left|X_n - \sum_{j=1}^{p} \alpha_j^\circ X_{n-j} - \mu_\varepsilon^\circ\right|^2 \middle| \mathcal{F}_{n-1}\right) \leq \sum_{n=1}^{\infty} \frac{1}{n^2}\left(\max_{1\leq k\leq p}\{\lambda_k\} \sum_{j=1}^{p} X_{n-j} + \sigma_\varepsilon^2\right)$$
$$= \max_{1\leq k\leq p}\{\lambda_k\} \sum_{n=1}^{\infty}\sum_{j=1}^{p} n^{-2} X_{n-j} + \frac{\pi^2\sigma_\varepsilon^2}{6}.$$

Puisque pour toute variable aléatoire X positive, $\mathbb{E}[X] < \infty$ implique $\Pr(X = \infty) = 0$, il suffit de montrer que

$$\mathbb{E}\left[\sum_{n=1}^{\infty}\sum_{j=1}^{p} n^{-2} X_{n-j}\right] < \infty$$

pour que la double sommation $\sum_{n=1}^{\infty}\sum_{j=1}^{p} n^{-2} X_{n-j}$ soit presque sûrement bornée. Or, par le théorème de la convergence monotone, nous avons

$$\mathbb{E}\left[\sum_{n=1}^{\infty}\sum_{j=1}^{p} n^{-2} X_{n-j}\right] = \lim_{m\to\infty}\sum_{n=1}^{m}\sum_{j=1}^{p} n^{-2}\mathbb{E}[X_{n-j}] = p\mu_X \frac{\pi^2}{6} < \infty.$$

Donc,

$$\sum_{n=1}^{\infty}\frac{1}{n^2}\mathbb{E}\left(\left|X_n - \sum_{j=1}^{p}\alpha_j^\circ X_{n-j} - \mu_\varepsilon^\circ\right|^2 \middle| \mathcal{F}_{n-1}\right) < \infty,\ \mathbb{P}_{\theta^\circ} - \text{p.s.}$$

En fait, nous venons de montrer que pour presque tout ω dans Ω,

$$\sum_{n=1}^{\infty}\frac{1}{n^2}\mathbb{E}\left(\left|X_n(\omega) - \sum_{j=1}^{p}\alpha_j^\circ X_{n-j}(\omega) - \mu_\varepsilon^\circ\right|^2 \middle| \mathcal{F}_{n-1}\right) < \infty$$
$$\text{et } \lim_{n\to\infty} U_n(\omega) = \lim_{n\to\infty} n = \infty.$$

Ainsi,

$$\mathbb{P}_{\theta^\circ}\left\{\omega : \lim_{n\to\infty} n = \infty \text{ et } \sum_{n=1}^{\infty}\frac{1}{n^2}\mathbb{E}\left(\left|X_n(\omega) - \sum_{j=1}^{p}\alpha_j^\circ X_{n-j}(\omega) - \mu_\varepsilon^\circ\right|^2 \middle| \mathcal{F}_{n-1}\right) < \infty\right\}$$
$$= 1. \quad \square$$

Preuve du lemme 5.4. Bornant presque sûrement

$$\mathbb{E}\left(\left|X_n - \sum_{j=1}^{p}\alpha_j^\circ X_{n-j} - \mu_\varepsilon^\circ\right|^2 \middle| \mathcal{F}_{n-1}\right)$$

par

$$\max_{1\le k\le p}\{\lambda_k\}\sum_{j=1}^{p} X_{n-j} + \sigma_\varepsilon^2$$

(voir la démonstration du lemme 5.3), l'expression

$$\sum_{n=1}^{\infty}\frac{1}{n^2}\mathbb{E}\left(\left|X_{n-i}\left(X_n-\sum_{j=1}^{p}\alpha_j^\circ X_{n-j}-\mu_\varepsilon^\circ\right)\right|^2\middle|\mathcal{F}_{n-1}\right)$$
$$=\sum_{n=1}^{\infty}\frac{1}{n^2}X_{n-i}^2\mathbb{E}\left(\left|X_n-\sum_{j=1}^{p}\alpha_j^\circ X_{n-j}-\mu_\varepsilon^\circ\right|^2\middle|\mathcal{F}_{n-1}\right)$$

est bornée presque sûrement par

$$\sum_{n=1}^{\infty}\frac{1}{n^2}X_{n-i}^2\left(\max_{1\le k\le p}\{\lambda_k\}\sum_{j=1}^{p}X_{n-j}+\sigma_\varepsilon^2\right)$$
$$=\max_{1\le k\le p}\{\lambda_k\}\sum_{n=1}^{\infty}\sum_{j=1}^{p}\frac{1}{n^2}X_{n-i}^2X_{n-j}+\sigma_\varepsilon^2\sum_{n=1}^{\infty}\frac{1}{n^2}X_{n-i}^2.$$

Comme $X_{n-j}\le X_{n-j}^2$ car X_{n-j} est une variable aléatoire à valeurs entières positives, la dernière expression est bornée par

$$\max_{1\le k\le p}\{\lambda_k\}\sum_{n=1}^{\infty}\sum_{j=1}^{p}\frac{1}{n^2}X_{n-i}^2X_{n-j}^2+\sigma_\varepsilon^2\sum_{n=1}^{\infty}\frac{1}{n^2}X_{n-i}^2. \tag{8.1}$$

On note que

$$\mathbb{E}\left(X_{n-i}^2X_{n-j}^2\right)=\mathrm{Cov}\left(X_{n-i}^2,X_{n-j}^2\right)+\mathbb{E}\left(X_{n-i}^2\right)\mathbb{E}\left(X_{n-j}^2\right)$$
$$=\varphi_0(i-j)+\left(\gamma(0)+\mu_X^2\right)^2<\infty$$

et le membre de droite de l'équation (8.1) est borné. La démonstration est complète. □

9. English extended abstract.

9.1. Introduction. The generalization of the STEUTEL and VAN HARN (1979) operator given by definition 1.1 is used to define the Generalized Integer-valued Autoregressive Process (GINAR) as stated in definition 3.1. Basic computational properties of this operator are given in section 2. Many proofs rely on probability generating functions. Let $\{X_t\}$ be a GINAR process. Proposition 3.1 gives $\mathbb{E}(X_t)$, the expected value of X_t, and Proposition 3.2 gives $\gamma(0)$, the variance of X_t. The links between the autocovariances (and the autocorrelations) and the parameters of the model are established in Proposition 3.3. The autocorrelation function satisfies the Yule-Walker equations. From this point of view the covariance structure of a GINAR process is identical to the structure of a standard gaussian process. The partial autocorrelation is defined in definition 3.2.

The estimation problem is tackled in section 4. First we consider the usual statistics used in time series analysis. Theorem 4.1 establishes that the sample mean is a strongly consistent estimator of $\mathbb{E}(X_t)$. Propositions 4.1, 4.2 and Theorem 4.2 give the properties of the sample autocovariances. All these classical estimators are asymptotically unbiased and, under some mild conditions, they are strongly consistent. The sample autocorrelations have similar properties by Theorem 4.3. More, the sample partial autocorrelations are strongly consistent.

In section 5, we recall some results of KLIMKO and NELSON (1978). These results are used to demonstrate that the conditional least squares estimators are strongly consistent. This is established in Theorem 5.3.

9.2. Generalized Steutel and van Harn operator and GINAR process.

Definition 2.1. (Generalized Steutel and van Harn operator). Let X be a non-negative integer-valued random variable; let Y be a non-negative integer-valued random variable with finite mean α and variance λ and let $\{Y_i\}_{i\in\mathbb{N}}$, be a sequence of i.i.d. random variables distributed as Y and independent of X. The operator $\alpha\circ$ is defined by $\alpha \circ X = \sum_{i=1}^{X} Y_i$.

The GINAR process is given in terms of this operator.

Definition 3.1. (GINAR process). Let $\{X_t\}_{t\in\mathbb{Z}}$ be a sequence of non-negative integer-valued random variables; let $\{\varepsilon_t\}_{t\in\mathbb{Z}}$, be a sequence of i.i.d. non-negative integer-valued random variables with finite mean μ_ε and finite variance σ_ε^2; let $p \in \mathbb{N}$ and $\{\alpha_k\}_{k\in\{1,2,\dots,p\}}$, a set of constants such that $\forall k \in \{1,2,\dots,p-1\}$, $\alpha_k \in [0,1)$, $\alpha_p \in (0,1)$ and $\sum_{k=1}^{p}\alpha_k < 1$. Then $\{X_t\}_{t\in\mathbb{Z}}$ is a GINAR(p) process if, and only if, $X_t = \sum_{k=1}^{p}\alpha_k \circ X_{t-k} + \varepsilon_t$. The sequences $\{Y_i^{(k)}\}_{i\in\mathbb{N}}$, ($k = 1,2,\dots,p$), of finite mean α_k and variance λ_k involved in $\alpha_k \circ X_{t-k}$ are independent amongst each other and are independent of ε_t.

Propositions 3.1 and 3.2 give the first two moments of X_t. The relation between the autocovariances (and the autocorrelations) and the parameters of the model is given in Proposition 3.3.

Proposition 3.1. *Let $\{X_t\}_{t\in\mathbb{Z}}$, be a GINAR(p) process. Then*

$$\mathbb{E}(X_t) = \mu_\varepsilon \left(1 - \sum_{k=1}^{p}\alpha_k\right)^{-1}.$$

Proposition 3.2. *Let $\{X_t\}_{t\in\mathbb{Z}}$, be a GINAR(p) process. The variance of X_t is*

$$\gamma(0) = \mu_X \sum_{i=1}^{p}\lambda_i + \sum_{i=1}^{p}\alpha_i\gamma(i) + \sigma_\varepsilon^2.$$

Proposition 3.3. *Let $\{X_t\}_{t\in\mathbb{Z}}$, be a GINAR(p) process. For $k \in \mathbb{N}$, $k \neq 0$, the autocovariance at lag k satisfies the relation $\gamma(k) = \sum_{i=1}^{p}\alpha_i\gamma(k-i)$, and the autocorrelation satisfies*

$$\rho(k) = \begin{cases} 1 & \text{if } k = 0 \\ \displaystyle\sum_{i=1}^{p}\alpha_i\rho(k-i) & \text{if } k \geq 1. \end{cases}$$

9.3. Classical statistics. The following assumptions are useful to demonstrate results about the strong consistency of the different estimators.

Assumptions. *Let $\gamma(\cdot)$ the autocovariance function of $\{X_t\}_{t\in\mathbb{Z}}$, a GINAR(p) process and let $\varphi_k(h) = Cov\,(X_tX_{t+k}, X_{t+h}X_{t+k+h})$. In the following, when necessary, the assumptions A1 and A2 will be required: A1: $\sum_{k=0}^{\infty}|\gamma(k)| < \infty$ and A2: $\sum_{h=0}^{\infty}|\varphi_k(h)| < \infty$.*

The classical statistics involved in time series analysis considered here are the sample mean ($\overline{X}$), the sample autocovariances and autocorrelations ($\hat{\gamma}(k)$ and $\hat{\rho}(k)$ respectively). Propositions 4.1 and 4.2 and Theorems 4.2 and 4.3 establish the strong consistency of these estimators.

Proposition 4.1. *For M fixed and $0 \leq k \leq M$, $\hat{\gamma}(k)$ is asymptotically unbiased for $\gamma(k)$ if assumption A1 is true.*

Proposition 4.2. *If A1 is true then for n fixed, the bias of $\hat{\gamma}(k)$ is of order n^{-1} and is, in general, increasing with k.*

Theorem 4.2. *Under assumptions A1 and A2, $\hat{\gamma}(k)$ is strongly consistent for $\gamma(k)$.*

Theorem 4.3. *Under the same assumptions, if $\gamma(0) > 0$, then $\hat{\rho}(k)$ is strongly consistent for $\rho(k)$.*

9.3. Conditional least squares estimators. The conditional least squares estimates of the parameters $\alpha_1, \alpha_2, \ldots, \alpha_p$ and μ_ε are obtained by minimizing the function $Q_n(\alpha, \mu_\varepsilon)$ of (5.1).

The main result concerning these estimators is given in Theorem 5.3.

Theorem 5.3. *If $\gamma(0) > 0$ and if assumption A2 holds, then, the solution of $\nabla Q_n(\alpha, \mu_\varepsilon) = 0$ generates strongly consistent estimators for $\alpha_1, \alpha_2, \ldots, \alpha_p$ and μ_ε.*

We conclude Section 5 by proving that the sample partial autocorrelations are also strongly consistent. Finally, in Section 8, we give the technical demonstrations that were not done in the main body of the text.

BIBLIOGRAPHIE

1. M. A. Al-Osh et A. A. Alzaid, *First order autoregressive time series with negative binomial and geometric marginals*, Comm. Statist. Theory Methods **21** (1992), 2483–2492.
2. ______, *Binomial Autoregressive Moving Average Models*, Comm. Statist. Stochastic Models **7** (1991), 261–282.
3. ______, *An integer-valued pth order autoregressive structure (INAR(p)) process*, J. Appl. Probab. **27** (1990), 314–324.
4. ______, *Integer-valued moving average (INMA) process*, Statist. Hefte **29** (1988), 281–300.
5. ______, *First-order integer-valued autoregressive (INAR(1)) process*, J. Time Ser. Anal. **8** (1987), 261–275.
6. T. W. Anderson, *The Statistical Analysis of Time Series*, Wiley, New York, 1971.
7. G. E. P. Box et G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, 1976.
8. D. R. Brillinger, *The calculation of cumulants via conditioning*, Ann. Inst. Statist. Math. **21** (1969).

9. P. J. Brockwell et R. A. Davis, *Time Series: Theory and Methods, 2^e^ éd.*, Springer-Verlag, New York, 1991.
10. J.-G. Du et Y. Li, *The integer-valued autoregressive (INAR(p)) model*, J. Time Ser. Anal. **12** (1991), 129–142.
11. G. Gauthier, *Modèles de type autorégressif pour les séries chronologiques à valeurs entières non négatives*, Mémoire de maîtrise, Université du Québec à Montréal, (1991).
12. G. R. Grimmett et D. R. Stirzaker, *Probability and random processes*, Clarendon Press, Oxford, 1983.
13. P. Hall et C. C. Heyde, *Martingale limit theory and its application*, Academic Press, New York, 1980.
14. L. A. Klimko et P. J. Nelson, *On conditional least squares estimation for stochastic processes*, Ann. Statist. **6** (1978).
15. P. A. Jacobs et P. A. W. Lewis, *Discrete time series generated by mixture I: correlational and runs properties*, J. Roy. Statist. Soc. Ser. B **40** (1978a), 94–105.
16. ———, *Discrete time series generated by mixture II: asymptotic properties*, J. Roy. Statist. Soc. Ser. B **40** (1978b), 222–228.
17. ———, *Discrete time series generated by mixture III: autoregressive processes (DAR(p))*, Rapp. de rech. no NPS55-78-022, Naval Postgraduate School, Monterey, California (1978c).
18. P. A. W. Lewis et P. A. Jacobs, *Stationary discrete autoregressive-moving average time series generated by mixtures*, J. Time Series Anal. **4** (1983), 19–36.
19. P. A. W. Lewis, *Simple models for positive-valued and discrete-valued time series with ARMA correlation structure*, North Holland, Amsterdam, 1980, pp. 151–166.
20. E. McKenzie, *Some ARMA models for dependent sequences of Poisson counts*, Adv. in Appl. Probab. **20** (1988), 822–835.
21. ———, *The distributional structure of finite moving-average processes*, J. Appl. Probab. **25** (1988), 313–321.
22. ———, *Autoregressive moving average processes with negative binomial and geometric marginal distributions*, Adv. in Appl. Probab. **18** (1986), 679–705.
23. ———, *Some simple models for discrete variate time series*, Wat. Res. Bul. **21** (1985), 645–650.
24. ———, *Extending the correlation structure of exponential autoregressive- moving-average processes*, J. Appl. Probab. **18** (1981), 181–189.
25. R. J. Serfling, *Approximation theorems of mathematical statistics*, Wiley, New York, 1980.
26. F. W. Steutel et K. van Harn, *Discrete analogues of self-decomposability and stability*, Ann. Probab. **7** (1979), 893–899.

G. GAUTHIER
CARLETON UNIVERSITY
DEPARTMENT OF MATHEMATICS AND STATISTICS
OTTAWA (ONTARIO) CANADA, K1S 5B6

A. LATOUR
UNIVERSITÉ DU QUÉBEC À MONTRÉAL
DÉPARTEMENT DE MATHÉMATIQUES ET D'INFORMATIQUE
CASE POSTALE 8888, SUCCURSALE CENTRE-VILLE
MONTRÉAL (QUÉBEC) CANADA, H3C 3P8