

La classe RINAR

Une nouvelle classe de modèles autorégressifs à valeurs entières

Présentée par

M. Kachour

Sous la direction du Professeur

J.F. Yao

IRMAR, Université de Rennes 1

9 Décembre 2009

Plan de la présentation

- 1 Motivation
- 2 INAR Vs. RINAR
- 3 Le modèle RINAR(1)
 - Stationnarité et ergodicité
 - Quelques propriétés
 - Estimation des paramètres
 - Variations du modèle
 - Exemple de corrélogramme
 - Méthode numérique pour calculer EMC
 - Une application réelle
- 4 Le modèle RINAR(p)
- 5 Le modèle RINVAR(1)
- 6 Perspectives
- 7 Quelques références

Plan de la présentation

- 1 Motivation
- 2 INAR Vs. RINAR
- 3 Le modèle RINAR(1)
 - Stationnarité et ergodicité
 - Quelques propriétés
 - Estimation des paramètres
 - Variations du modèle
 - Exemple de corrélogramme
 - Méthode numérique pour calculer EMC
 - Une application réelle
- 4 Le modèle RINAR(p)
- 5 Le modèle RINVAR(1)
- 6 Perspectives
- 7 Quelques références

Exemples :

- Nombre de cas d'une certaine maladie
- Nombre des transactions quotidiennes d'un titre boursier
- Nombre mensuel d'accidents mortels sur une autoroute
- Température quotidienne à Rennes pendant cet été
- Et bien d'autres ...

Domaines d'application :

- Médecine (épidémiologie)
- Finance et économie
- Assurance et fiabilité
- Météorologie
- Et bien d'autres ...

Polio, États-Unis

Données mensuelles entre Janvier 1970 et Décembre 1983

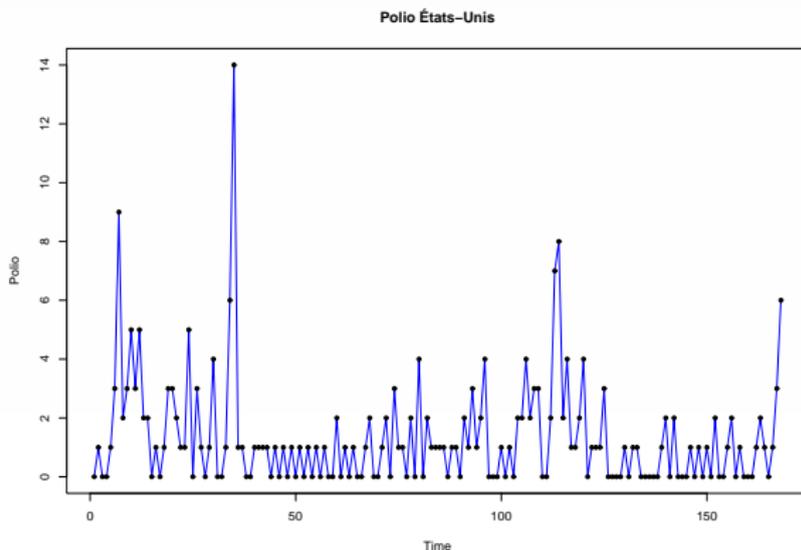


Figure: 168 observations, source : [Brockwell et Davis (1991)]

Plan de la présentation

- 1 Motivation
- 2 INAR Vs. RINAR**
- 3 Le modèle RINAR(1)
 - Stationnarité et ergodicité
 - Quelques propriétés
 - Estimation des paramètres
 - Variations du modèle
 - Exemple de corrélogramme
 - Méthode numérique pour calculer EMC
 - Une application réelle
- 4 Le modèle RINAR(p)
- 5 Le modèle RINVAR(1)
- 6 Perspectives
- 7 Quelques références

L'opérateur de Steutel et Van Harn (1979) :

Soit $Z \in \mathbb{N}$, une variable aléatoire; $a \in [0, 1]$

$$a \circ Z = \sum_{k=1}^Z \xi_k,$$

où (ξ_k) suite i.i.d. Bernoulli(a), indépendante de Z , série de comptage

Le modèle INAR(p):

$$X_t = \sum_{j=1}^p a_j \circ X_{t-j} + \varepsilon_t,$$

où (ε_t) i.i.d. \mathbb{N} , indépendantes de toutes les séries de comptage

Propriétés des INAR

- $INAR(1)$ fonction d'autocorrélation identique à celle du $AR(1)$ réel (cf. Al-osh et Alzaid [2], 1987)
- Pour $p > 1$, deux différentes spécifications :
 - $INAR(p) - AA$ fonction d'autocorrélation similaire à celle du $ARMA(p, p - 1)$ réel (cf. Al-osh et Alzaid [3], 1990)
 - $INAR(p) - DL$ fonction d'autocorrélation identique à celle du $AR(p)$ réel (cf. Du et Li [4], 1991)

Limitations des INAR :

- Une structure d'**innovation forte** complexe : (ε_t) et séries de comptage
- Les *INAR* ne peuvent pas produire des autocorrélations négatives
- Les *INAR* ne peuvent pas modéliser des séries chronologiques à valeurs entières avec des observations négatives

Le modèle RINAR(p) :

$$X_t = \left\langle \sum_{j=1}^p \alpha_j X_{t-j} + \lambda \right\rangle + \varepsilon_t,$$

où

- $\langle x \rangle$ l'arrondi de $x \in \mathbb{R}$ à l'entier le plus près,
- (ε_t) i.i.d. \mathbb{Z} , $\mathbb{E}[\varepsilon_t] = 0$

Remarques :

- λ paramètre de décentrage
- $\langle \cdot \rangle$ opération naturelle pour passer des nombres réels aux entiers

Avantages des RINAR :

- Une structure d'innovation forte simple : (ε_t) seul
- Les *RINAR* peuvent produire des autocorrélations négatives
- Les *RINAR* peuvent modéliser des séries chronologiques avec des valeurs négatives
- La prévision à un pas, basée sur l'espérance conditionnelle, est entière et donnée par

$$\hat{X}_T = \mathbb{E}(X_T | X_s, s \leq T-1) = \left\langle \sum_{j=1}^p \alpha_j X_{T-j} + \lambda \right\rangle$$

Plan de la présentation

- 1 Motivation
- 2 INAR Vs. RINAR
- 3 Le modèle RINAR(1)**
 - Stationnarité et ergodicité
 - Quelques propriétés
 - Estimation des paramètres
 - Variations du modèle
 - Exemple de corrélogramme
 - Méthode numérique pour calculer EMC
 - Une application réelle
- 4 Le modèle RINAR(p)
- 5 Le modèle RINVAR(1)
- 6 Perspectives
- 7 Quelques références

- RINAR(1) est défini par

$$X_t = \langle \alpha X_{t-1} + \lambda \rangle + \varepsilon_t = f_\theta(X_{t-1}) + \varepsilon_t, \quad t \in \mathbb{Z}$$

- Fonction de régression $f_\theta(x) = \langle \alpha x + \lambda \rangle$, $x \in \mathbb{Z}$ et $\theta = (\alpha, \lambda) \in \Theta$
- Comme (ε_t) est i.i.d., (X_t) est une chaîne de Markov homogène sur $E = \mathbb{Z}$ avec noyau de transition

$$\pi(x, y) = \mathbb{P} \{ \varepsilon_1 = y - f_\theta(x) \}, \quad x, y \in E$$

Proposition :

Soit $\theta = (\alpha, \lambda) \in \Theta$ fixé. Supposons que :

- 1 (X_t) est irréductible
- 2 $\mathbb{E}|\varepsilon_t|^k < +\infty$, pour un certain $k \geq 1$
- 3 $|\alpha| < 1$

Alors,

- (X_t) possède une unique loi invariante, notée μ . De plus, μ possède un moment d'ordre k
- Pour tout $x \in E$ et $f \in L^1(\mu)$ nous avons

$$\frac{1}{n} \sum_{k=1}^n f(X_k) \longrightarrow \mu(f), \quad \mathbb{P}_x \text{ p.s.}$$

où \mathbb{P}_x représente la probabilité conditionnelle $\mathbb{P}(\cdot \mid X_0 = x)$

Résumé de la preuve :

- $V : x \mapsto |x|^k$, où $k \geq 1$, fonction de Lyapunov
- Comme $|\alpha| < 1$ et $\mathbb{E}|\varepsilon_t|^k < +\infty$, nous obtenons

$$\limsup_{|x| \rightarrow \infty} \frac{\pi V(x)}{V(x)} < 1 \Leftrightarrow \pi V(x) \leq \gamma V(x) + \delta,$$

où $0 \leq \gamma < 1$ et $\delta < \infty$.

- $\limsup_n \mu_n V < \infty$, où μ_n est la mesure empirique de (X_t)
- (μ_n) est une suite tendue et pour tout point limite μ nous avons $\mu V < \infty$
- (X_t) est irréductible, μ est unique

Soit (X_t) le processus RINAR(1) stationnaire. Nous notons,

$$m = \mathbb{E}X_t \text{ et } \rho(k) = \frac{\text{cov}(X_t, X_{t+k})}{V(X_t)}, \quad \forall t \in \mathbb{Z}, k \in \mathbb{N}$$

Proposition :

$$\left| m - \frac{\lambda}{1 - \alpha} \right| \leq \frac{1}{2(1 - \alpha)}$$

Ainsi, nous avons

$$m = 0 \Rightarrow |\lambda| \leq \frac{1}{2}, \quad \lambda = 0 \Rightarrow |m| \leq \frac{1}{2(1 - \alpha)}$$

$$\lambda \in \mathbb{Z} \Rightarrow \lambda = \langle m(1 - \alpha) \rangle$$

Proposition :

Si $m \in \mathbb{Z}$, alors

$$|\rho(k) - \alpha^k| \leq \frac{1}{2} \sum_{i=0}^{k-1} |\alpha|^i$$

Si $m \notin \mathbb{Z}$, alors

$$|\rho(k) - \alpha^k| \leq \frac{1}{2} \left(1 + \frac{\mu([m])}{\Sigma'} \right) \sum_{i=0}^{k-1} |\alpha|^i,$$

où μ est la loi stationnaire de (X_t) , $[\cdot]$ partie entière,

$$\Sigma' = \sum_{i \neq [m]} \mu(i) |i - m| = \mathbb{E}[|X_t - m|] - \mu([m]) |[m] - m|$$

Quelques notations

- $\theta_0 = (\alpha_0, \lambda_0)$ la vraie valeur du paramètre θ
- \mathbb{P}_{θ_0} la distribution de probabilité de (X_t) sous le vrai modèle
- Toute convergence $\xrightarrow{p.s.}$ signifie une convergence presque sûre sous \mathbb{P}_{θ_0}

Hypothèse [H1]

- 1 Sous \mathbb{P}_{θ_0} , (X_t) est irréductible
- 2 pour un certain $k \geq 2$, $\mathbb{E}|\varepsilon_t|^k < +\infty$
- 3 $\Theta \subset]-1, 1[\times \mathbb{R}$, compact

Conséquences

- (X_t) possède une unique loi invariante μ_{θ_0} telle que $\mu_{\theta_0}(|\cdot|^k) < \infty$ avec $k \geq 2$
- $Y_t = (X_{t-1}, X_t)$ possède aussi une unique loi invariante P telle que $P(\|\cdot\|^k) < \infty$ pour toute norme $\|\cdot\|$ sur $E^2 = \mathbb{Z}^2$

Soient X_0, X_1, \dots, X_n des observations du processus RINAR(1).
L'estimateur des moindres carrés est défini par

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \frac{1}{n} \sum_{t=1}^n [X_t - f_{\theta}(X_{t-1})]^2 = \arg \min_{\theta \in \Theta} P_n g_{\theta},$$

P_n est la mesure empirique générée par Y_1, \dots, Y_n

$$g_{\theta}(z) = [y - f_{\theta}(x)]^2, \quad \forall z = (x, y) \in E^2 \text{ et } \theta \in \Theta$$

Problème d'identifiabilité :

Soient $\theta = (\alpha, \lambda) \in \Theta$ et $\{x\}$ la partie fractionnaire de $x \in \mathbb{R}$

$$f_\theta = f_{\theta_0}, \mu_{\theta_0} - p.s. \Leftrightarrow \langle \alpha x + \lambda \rangle = \langle \alpha_0 x + \lambda_0 \rangle, \forall x \in \mathbb{Z}$$

$$\Rightarrow \alpha = \alpha_0, \langle \lambda \rangle = \langle \lambda_0 \rangle$$

- Si $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$, alors $(\{\alpha_0 x\})_{x \in \mathbb{Z}}$ dense dans $[0, 1[$ et par suite

$$\lambda = \lambda_0$$

- Si $\alpha_0 = \frac{p}{q}$, où $p \in \mathbb{Z}$ $q \in \mathbb{N}^*$, $p \wedge q = 1$, alors

$$\forall x \in \mathbb{Z}, \{\alpha_0 x\} \in \left\{ 0, \frac{1}{q}, \frac{2}{q}, \dots, \frac{q-1}{q} \right\} \text{ et par suite}$$

$$\lambda \in I_0 \ni \lambda_0, |I_0| = \frac{1}{q}$$

Problème d'identifiabilité :

Exemple : supposons que $\alpha = \alpha_0 = \frac{1}{4}$ et $\lambda \geq \lambda_0 = 2.2$

Ici, $q = 4$ et $\{\lambda_0\} = 0.2 \in i_0 = \left[0, \frac{1}{4}\right[$ ($\langle \lambda \rangle = \langle \lambda_0 \rangle = 2$)

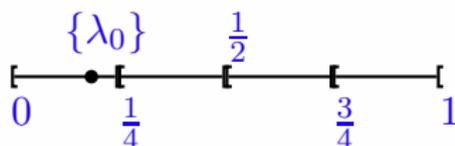


Figure: La subdivision de l'intervalle $[0, 1[$ en 4 intervalles de taille $\frac{1}{4}$

$$f_\theta = f_{\theta_0}, \mu_{\theta_0} - p.s. \Rightarrow \langle \{\alpha_0 x\} + \{\lambda\} \rangle = \langle \{\alpha_0 x\} + \{\lambda_0\} \rangle, \forall x \in \mathbb{Z}$$

$$I_0 = \{\lambda, \langle \lambda \rangle = \langle \lambda_0 \rangle \text{ et } \{\lambda\} \in i_0\} = [2; 2.25[, \quad |I_0| = 0.25 = \frac{1}{4}$$

Soit d distance sur Θ , $\forall \theta = (\alpha, \lambda), \theta' = (\alpha', \lambda') \in \Theta$,

$$d(\theta, \theta') = \max \{ |\alpha - \alpha'|, |\lambda - \lambda'| \}$$

Théorème :

Supposons que [H1] soit vérifiée.

- Si $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$, alors

$$\hat{\theta}_n \xrightarrow{p.s.} \theta_0$$

- Si $\alpha_0 = \frac{p}{q}$, où $p \in \mathbb{Z}$ $q \in \mathbb{N}^*$, $p \wedge q = 1$, alors

$$d(\hat{\theta}_n, A_0) \xrightarrow{p.s.} 0, \text{ où } A_0 = \{\alpha_0\} \times I_0$$

Étapes de la preuve

- LGNU : $\sup_{\theta \in \Theta} |(P_n - P)g_\theta| \xrightarrow{p.s.} 0$
 - $\forall \theta \in \Theta, y \in \mathbb{Z}^2, g_\theta(y) \leq A(y) = c(1 + \|y\|^2) \in L^1(P)$
 - Soit $q > 0$ fixé,

$$\begin{aligned} \sup_{\theta \in \Theta} |(P_n - P)g_\theta| &\leq \sum_{\|y\| < q} A(y) |P_n(y) - P(y)| \\ &\quad + (P_n + P)(A(y) \mathbb{1}_{\|y\| \geq q}) \end{aligned}$$

- Comme $P_n(y) \xrightarrow{p.s.} P(y)$, nous avons

$$\limsup_n \sup_{\theta \in \Theta} |(P_n - P)g_\theta| \leq 2P(A(y) \mathbb{1}_{\|y\| \geq q}), \text{ p.s.}$$

- En prenant $q \nearrow \infty$, nous obtenons

$$\limsup_n \sup_{\theta \in \Theta} |(P_n - P)g_\theta| = 0, \text{ p.s.}$$

Étapes de la preuve

- Bonne discrimination du contraste limite :
Pour tout (suffisamment petit) $\varepsilon > 0$,
 - si $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$, alors

$$\inf_{\theta \in \Theta_\varepsilon} |P(g_\theta - g_{\theta_0})| > 0, \text{ où } \Theta_\varepsilon = \{\theta : d(\theta, \theta_0) \geq \varepsilon\}$$

- si $\alpha_0 = \frac{p}{q}$, où $p \in \mathbb{Z}$ $q \in \mathbb{N}^*$, $p \wedge q = 1$, alors

$$\inf_{\theta \in \Theta_\varepsilon^0} |P(g_\theta - g_{\theta_0})| > 0, \text{ où } \Theta_\varepsilon^0 = \{\theta : d(\theta, A_0) \geq \varepsilon\}$$

- Arguments standards de la théorie des M-estimateurs

- Un modèle totalemtent identifiable :

$$X_t = \langle \alpha X_{t-1} \rangle + \varepsilon_t,$$

où (ε_t) i.i.d. \mathbb{Z} , $\mathbb{E}[\varepsilon_t] = \lambda$

Théorème :

Supposons que [H1] soit vérifiée. Alors, $\hat{\theta}_n \xrightarrow{p.s.} \theta_0$

Cas intéressant : $\lambda \in \mathbb{Z}$ (en particulier, $\lambda = 0$)

- Si $\Theta \subset [0, 1[\times]0, \infty[$ compact et (ε_t) i.i.d. \mathbb{N} , nous obtenons un modèle à valeurs entières positives, le processus PRINAR(1)

- Nous considérons le modèle suivant

$$X_t = \langle \alpha X_{t-1} \rangle + \varepsilon_t,$$

(ε_t) i.i.d. centrées dans \mathbb{Z} , pour $k \geq 2$, $\mathbb{E}|\varepsilon_t|^k < \infty$,
 $\alpha \in \Theta \subset]-1, 1[$ compact et (X_t) est irréductible

Hypothèse [A]

- 1 ε_t possède une loi symétrique sur $E = \mathbb{Z}$
- 2 $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$

$$\mathbb{E}[X_t] = 0, \forall t \in \mathbb{Z}$$

Ce modèle est appelé RINAR(1) centré

Nous considérons le modèle RINAR(1) centré $X_t = \langle \alpha X_{t-1} \rangle + \varepsilon_t$

Conséquences

- $m = \mathbb{E}[X_t] = 0 \in \mathbb{Z}$
- $\forall k \in \mathbb{N}^*, |\rho(k) - \alpha^k| \leq \frac{1}{2} \sum_{i=0}^{k-1} |\alpha|^i$

Nous simulons 1000 observations avec :

$$\alpha_0 = 0.83 + \frac{\pi}{800}, \varepsilon_t = A_t - B_t, \text{ où } A_t, B_t, \text{ i.i.d. } Po(\nu)$$

- Grand bruit, $\nu \gg 1$
- Petit bruit, $0 < \nu \leq \frac{1}{2}$

- Cas 1 : $\alpha_0 = 0.83 + \frac{\pi}{800}$ et $\nu = 9.453$

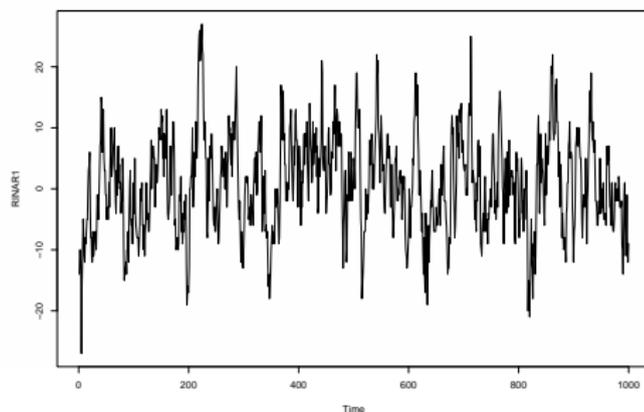


Figure: Les observations simulées

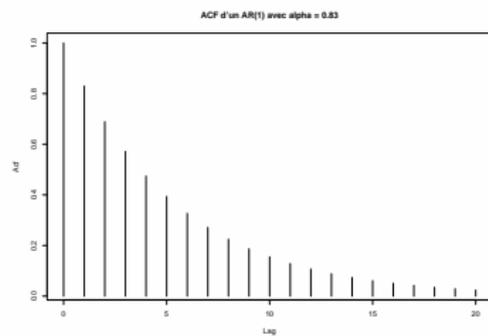
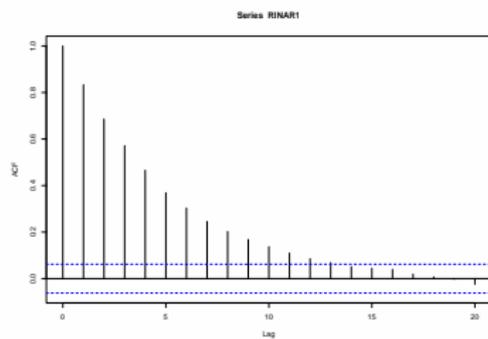


Figure: (1) ACF empirique de la série simulée (2) ACF théorique AR(1) avec $\alpha = 0.83$

- Cas 2 : $\alpha_0 = 0.83 + \frac{\pi}{800}$ et $\nu = 0.127$

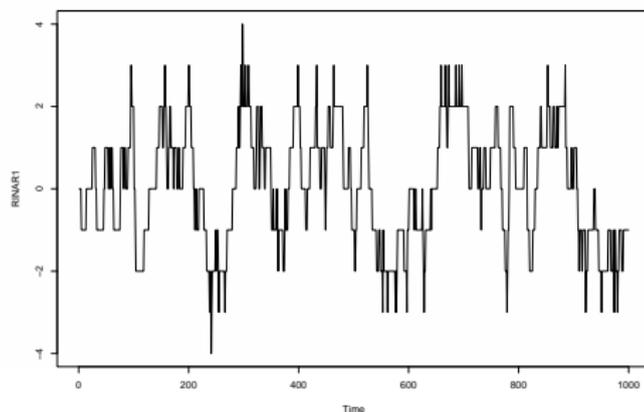


Figure: Les observations simulées

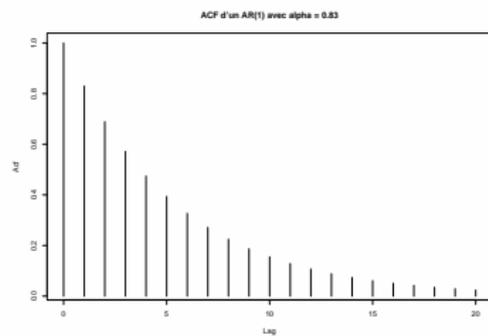
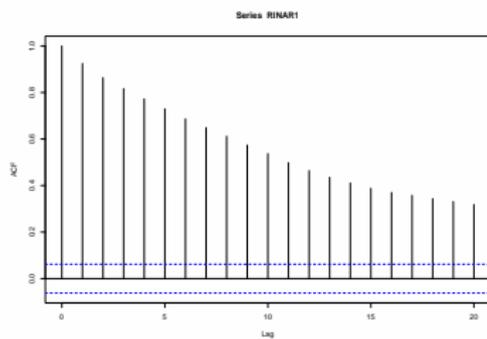


Figure: (1) ACF empirique de la série simulée (2) ACF théorique AR(1) avec $\alpha = 0.83$

X_0, X_1, \dots, X_n des observations du processus RINAR(1)

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \frac{1}{n} \sum_{t=1}^n [X_t - \langle \alpha X_{t-1} + \lambda \rangle]^2$$

- 1 Point de départ : $\hat{\theta}_0 = (\hat{\alpha}_0 = \hat{\rho}(1), \hat{\lambda}_0 = \bar{X}_n(1 - \hat{\alpha}_0))$
- 2 Transition de $\hat{\theta}_k = (\hat{\alpha}_k, \hat{\lambda}_k)$ à $\hat{\theta}_{k+1} = (\hat{\alpha}_{k+1}, \hat{\lambda}_{k+1})$
 - $\hat{\lambda}_k$ fixé, recherche dichotomique successive sur α , avec $] -1, 1[$ intervalle de recherche
 - $\hat{\alpha}_{k+1}$ fixé, recherche dichotomique successive sur λ , avec $[\hat{\lambda}_0 - 5 |\hat{\lambda}_0|, \hat{\lambda}_0 + 5 |\hat{\lambda}_0|]$ intervalle de recherche
- 3 Critère d'arrêt : $\max \left\{ |\hat{\alpha}_{k+1} - \hat{\alpha}_k|, |\hat{\lambda}_{k+1} - \hat{\lambda}_k| \right\} \leq 0.001$

Validation de l'algorithme

- Simulation de 500 observations du processus RINAR(1),
avec $\theta_0 = (\alpha_0, \lambda_0)$ fixé et $\varepsilon_t = A_t - B_t$
où A_t, B_t , i.i.d. $Po(\nu)$, avec $\nu \gg 1$
- Calcul de $\hat{\theta}_n$ avec l'algorithme
- Après 500 répétitions, nous obtenons $\{\hat{\alpha}_{n,i}, 1 \leq i \leq 500\}$ et $\{\hat{\lambda}_{n,i}, 1 \leq i \leq 500\}$

Validation de l'algorithme

Exemple : $\alpha_0 = -0.4$, $\lambda_0 = 1.44$ et $\nu = 6.23$

$\tilde{\alpha}_n = -0.4075$ (*s.e.* 0.045) et $\tilde{\lambda}_n = 1.3988$ (*s.e.* 0.162)

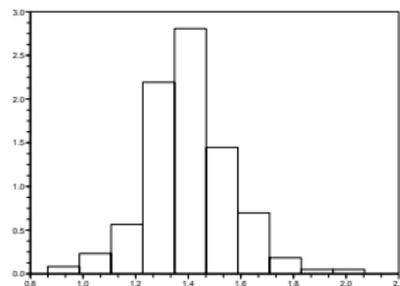
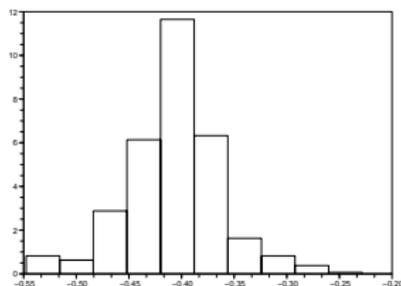


Figure: (1) Histogrammes des $(\hat{\alpha}_{n,i})_{1 \leq i \leq 500}$ (2) Histogrammes des $(\hat{\lambda}_{n,i})_{1 \leq i \leq 500}$

Analyse des taux de variation annuels de la population Suédoise 1750-1849.

Source : [McCleary et Hay, 1980]

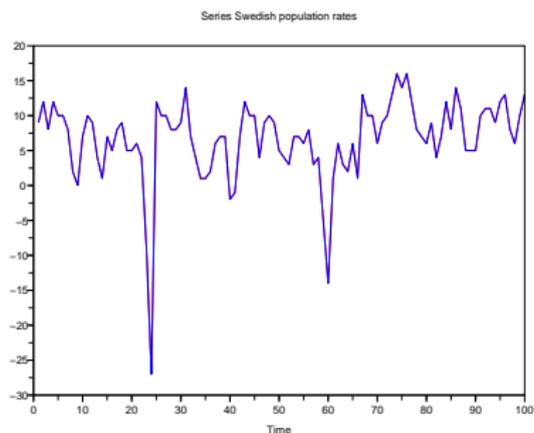


Figure: Les taux de variation annuels de la population Suédoise, 1750-1849

Analyse des taux de variation annuels de la population Suédoise 1750-1849

80 premières observations pour l'estimation et 20 dernières ensemble test pour la prévision

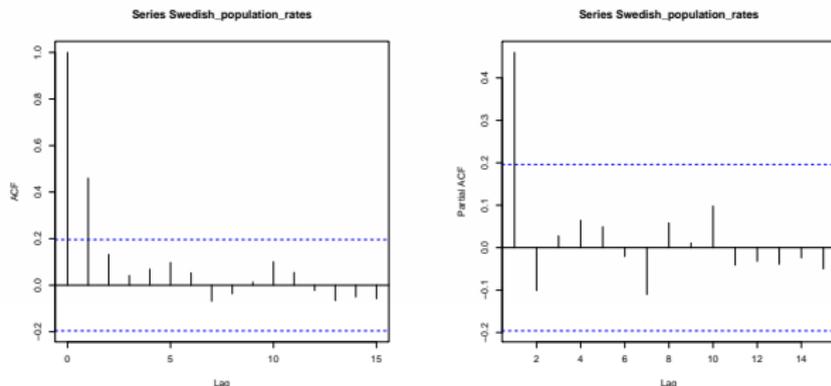


Figure: ACF et PACF empiriques

Analyse des taux de variation annuels de la population Suédoise 1750-1849

Ajustement à un AR(1) réel :

$$X_t = aX_{t-1} + b + \eta_t$$

$$\hat{a} = \hat{\rho}(1) = 0.4505, \quad \hat{b} = \bar{X}_n(1 - \hat{a}) = 3.337$$

Prévision à un pas (suivie d'arrondi) : $\tilde{X}_{T+1} = \langle \hat{a}X_T + \hat{b} \rangle$

$$EQM = \frac{1}{20} \sum_{T=80}^{99} \left(X_{T+1} - \tilde{X}_{T+1} \right)^2 = 5.8$$

et

$$EAM = \frac{1}{20} \sum_{T=80}^{99} |X_{T+1} - \tilde{X}_{T+1}| = 2.1$$

Analyse des taux de variation annuels de la population Suédoise 1750-1849

Ajustement à un RINAR(1) :

$$X_t = \langle \alpha X_{t-1} + \lambda \rangle + \varepsilon_t$$

$$\hat{\theta}_n = (\hat{\alpha}_n, \hat{\lambda}_n) = (0.469, 3.559)$$

Prévision à un pas : $\hat{X}_{T+1} = \langle \hat{\alpha}_n X_T + \hat{\lambda}_n \rangle$

$$EQM = \frac{1}{20} \sum_{T=80}^{99} \left(X_{T+1} - \hat{X}_{T+1} \right)^2 = 4$$

et

$$EAM = \frac{1}{20} \sum_{T=80}^{99} |X_{T+1} - \hat{X}_{T+1}| = 1.7$$

Analyse des taux de variation annuels de la population Suédoise 1750-1849

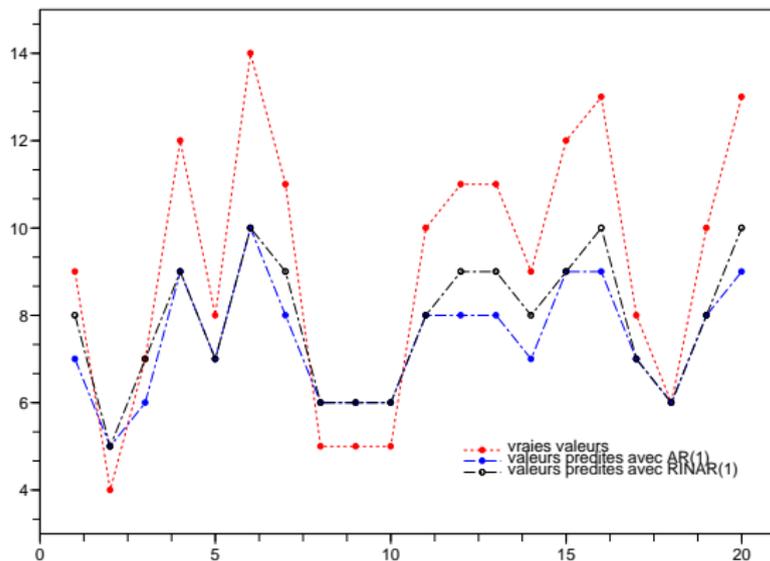


Figure: Les 20 dernières observations et leurs prévisions basées sur RINAR(1) et AR(1)

Plan de la présentation

- 1 Motivation
- 2 INAR Vs. RINAR
- 3 Le modèle RINAR(1)
 - Stationnarité et ergodicité
 - Quelques propriétés
 - Estimation des paramètres
 - Variations du modèle
 - Exemple de corrélogramme
 - Méthode numérique pour calculer EMC
 - Une application réelle
- 4 Le modèle RINAR(p)
- 5 Le modèle RINVAR(1)
- 6 Perspectives
- 7 Quelques références

Nous considérons le modèle RINAR(2) :

$$X_t = \langle \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \lambda \rangle + \varepsilon_t, \quad t \in \mathbb{Z}$$

Fonction de regression $f_\theta(z) = \langle \alpha_1 x + \alpha_2 y + \lambda \rangle$,
où $z = (x, y) \in \mathbb{Z}^2$ et $\theta = (\alpha_1, \alpha_2, \lambda)$

Conditions de stationnarité et d'ergodicité :

- (X_t, X_{t-1}) est irréductible
- (ε_t) possède un moment d'ordre $k \geq 2$
- $|\alpha_1| + |\alpha_2| < 1$

$\theta = (\alpha_1, \alpha_2, \lambda) \in \Theta$, compact sous-ensemble de $] -1, 1[^2 \times \mathbb{R}$

- 1 $\theta_0 = (\alpha_{1,0}; \alpha_{2,0}; \lambda_0)$ vraies valeurs des paramètres du modèle
- 2 $\hat{\theta}_n$ l'estimateur des moindres carrés
- 3 L'identifiabilité du modèle a un comportement non-standard

- Si $\alpha_{1,0} = \frac{a_1}{b_1}$ et $\alpha_{2,0} = \frac{a_2}{b_2}$, où $a_1 \wedge b_1 = 1$ et $a_2 \wedge b_2 = 1$, alors

$$\hat{\theta}_n \xrightarrow{p.s.} E_0, \quad E_0 = \{(\alpha_{1,0}; \alpha_{2,0})\} \times I_0 \quad (\lambda_0 \in I_0)$$

De plus,

$$\frac{1}{b_1 b_2} \leq |I_0| \leq \min \left\{ \frac{1}{b_1}, \frac{1}{b_2} \right\}$$

- Si $\alpha_{1,0}$ (ou $\alpha_{2,0}$) est irrationnel, alors

$$\hat{\theta}_n \xrightarrow{p.s.} \theta_0$$

Plan de la présentation

- 1 Motivation
- 2 INAR Vs. RINAR
- 3 Le modèle RINAR(1)
 - Stationnarité et ergodicité
 - Quelques propriétés
 - Estimation des paramètres
 - Variations du modèle
 - Exemple de corrélogramme
 - Méthode numérique pour calculer EMC
 - Une application réelle
- 4 Le modèle RINAR(p)
- 5 Le modèle RINVAR(1)**
- 6 Perspectives
- 7 Quelques références

Soit $X_t = \begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix}$. Nous considérons le modèle RINVAR(1) :

$$X_t = \langle MX_{t-1} + \lambda \rangle + \varepsilon_t = F_\theta(X_{t-1}) + \varepsilon_t,$$

où

$$\theta = (M, \lambda), M = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}, \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}, \varepsilon_t = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}$$

Par suite,

$$\begin{cases} X_{1,t} = \langle \alpha_{11}X_{1,t-1} + \alpha_{12}X_{2,t-1} + \lambda_1 \rangle + \varepsilon_{1,t} \\ X_{2,t} = \langle \alpha_{21}X_{1,t-1} + \alpha_{22}X_{2,t-1} + \lambda_2 \rangle + \varepsilon_{2,t} \end{cases}$$

$$\theta_0 = (M_0, \lambda_0) \text{ vraie valeur, } M_0 = (\alpha_{i,j}^0) \text{ et } \lambda_0 = \begin{pmatrix} \lambda_1^0 \\ \lambda_2^0 \end{pmatrix}$$

Conditions de stationnarité et d'ergodicité :

- (X_t) est irréductible
- (ε_t) possède un moment d'ordre $k \geq 2$
- $\rho(A) < 1$, où $A = |M| = \begin{pmatrix} |\alpha_{11}| & |\alpha_{12}| \\ |\alpha_{21}| & |\alpha_{22}| \end{pmatrix}$

L'identifiabilité du modèle :

- Chaque ligne de la matrice M_0 possède au moins un coefficient irrationnel ($\hat{\theta}_n \xrightarrow{p.s.} \theta_0$)
- Tous les coefficients de M_0 sont rationnels
- Une ligne de M_0 a au moins un coefficient irrationnel et l'autre ne possède que des coefficients rationnels

$$\left(\hat{\theta}_n \xrightarrow{p.s.} M_0 \times \begin{pmatrix} I_{0,1} \\ I_{0,2} \end{pmatrix} \right)$$

Plan de la présentation

- 1 Motivation
- 2 INAR Vs. RINAR
- 3 Le modèle RINAR(1)
 - Stationnarité et ergodicité
 - Quelques propriétés
 - Estimation des paramètres
 - Variations du modèle
 - Exemple de corrélogramme
 - Méthode numérique pour calculer EMC
 - Une application réelle
- 4 Le modèle RINAR(p)
- 5 Le modèle RINVAR(1)
- 6 Perspectives
- 7 Quelques références

L'étude de la vitesse de convergence de l'estimateur des moindres carrés, $\hat{\theta}_n$, pour le processus RINAR(1)

Travail en cours avec **Adrien Saumard**

$$g_{\hat{\theta}_n} = \arg \min_{g_{\theta} \in \mathcal{G}} P_n g_{\theta},$$

$\mathcal{G} = (g_{\theta})_{\theta \in \Theta}$, où $g_{\theta}((x, y)) = (y - f_{\theta}(x))^2$, $(x, y) \in \mathbb{Z}^2$

- (X_t) est β -mélangeante
- \mathcal{G} est une VC-classe de sous-graphes
-

$$\mathbb{E}_{\mu_{\theta_0}} (f_{\hat{\theta}_n} - f_{\theta_0})^2 \leq \frac{1}{\sqrt{n}} \mathbb{E} \left[\sup_{g_{\theta} \in \mathcal{G}} |\alpha_n (g_{\theta} - g_{\theta_0})| \right],$$

où $\alpha_n = \sqrt{n} (P - P_n)$

- Contrôle du supremum du processus empirique par entropie universelle pour les données mélangeantes

- Trouver la loi asymptotique de l'estimateur des moindres carrés
- Examiner l'innovation linéaire du processus RINAR(1)
- Étudier la stationnarité et d'ergodicité sous l'hypothèse d'un bruit blanc
- Améliorer les conditions de stationnarité et d'ergodicité pour RINAR(p) et RINVAR(1)
- Expliquer théoriquement le comportement du corrélogramme du RINAR(1) en fonction du choix du bruit (grand ou petit)
- Combiner le principe d'arrondi et les modèles ARCH/GARCH pour des séries chronologiques à valeurs entières

Plan de la présentation

- 1 Motivation
- 2 INAR Vs. RINAR
- 3 Le modèle RINAR(1)
 - Stationnarité et ergodicité
 - Quelques propriétés
 - Estimation des paramètres
 - Variations du modèle
 - Exemple de corrélogramme
 - Méthode numérique pour calculer EMC
 - Une application réelle
- 4 Le modèle RINAR(p)
- 5 Le modèle RINVAR(1)
- 6 Perspectives
- 7 Quelques références



McKenzie, E. (1985). Some simple models for discrete variate time series. Water Resources Bulletin, Volume 21, P. 645-650.



Alzaid, A. A. and Al-Osh, M. (1987). First-order integer-valued autoregressive (INAR(1)) process. J. Time Ser. Anal., volume 8, P. 261-275.



Alzaid, A. A. and Al-Osh, M. (1990). An integer-valued p th-order autoregressive structure (INAR(p)) process. J. Appl. Probab., volume 27, P. 314-324.



Du, J.-G. and Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model. J. Times Ser. Anal., Volume 12, P. 129-142.



Kachour, M. and Yao, J.F. (2009). First-order rounded integer-valued autoregressive (RINAR(1)) process. J. Times Ser. Anal., Volume 4, P. 417-448.



Kachour, M. (2009). p -order rounded integer-valued autoregressive (RINAR(p)) process. Preprint, University of Rennes 1.



Kachour, M. (2009). First-order rounded integer-valued vectorial autoregressive (RINVAR(1)) process. Preprint, University of Rennes 1.