Convergence des moments dans le théorème de la limite centrale presque sûr pour les martingales vectorielles

Bernard Bercu*, Peggy Cénac†, Guy Fayolle ‡

Thème BIO — Systèmes biologiques

Projet Preval

Rapport de recherche n° 6056 — November 2006 — 21 pages

Résumé : On établit dans ce rapport de nouvelles propriétés de convergence presque sûre de transformées de martingales vectorielles. On montre en particulier que, sous certaines conditions de régularité du processus croissant et sous certaines conditions de moments sur la martingale, il y a convergence des moments normalisés de tout ordre pair dans le théorème de la limite centrale presque sûr pour les martingales vectorielles. On formule également une conjecture de convergence sous des hypothèses moins restrictives, couvrant des familles plus vastes de processus. Enfin, on applique ces résultats aux modèles de régression linéaire ainsi qu'à certains processus de branchement avec immigration, ce qui permet d'établir de nouvelles propriétés asymptotiques sur les erreurs d'estimation et de prédiction.

Mots-clés : Martingale vectorielle, moment, régression, transformée, théorème central limite presque sûr.

* Institut de mathématiques, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence cedex

† MAP 5 Mathématiques Appliquées Université Paris 5, UMR CNRS 8145, 45 rue des Saints-Pères, 75270 Paris cedex 06

‡ INRIA Domaine de Voluceau B.P.105 78 153 Le Chesnay Cedex (France)

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 —Télécopie : +33 1 39 63 53 30

Convergence of moments in the almost sure central Limit theorem for multivariate martingales

Abstract: In this report, new almost sure convergence properties for multivariate martingale transforms are established. Assuming some regularity conditions both on the increasing process and on the moments of the martingale, one can show that normalized moments of any even order do also converge in the almost sure cental limit theorem for martingales. In addition, a conjecture about convergence under wider hypotheses is formulated, the goal being to cover a larger class of processes. Cases studies are proposed, including linear models and some branching processes with immigration, for which new asymptotic estimation and prediction errors are obtained.

Key-words: Almost sure cental limit theorem, moment, regression, multivariate martingale.

1 Introduction

Soit (X_n) une suite de variables aléatoires réelles indépendantes et de même loi avec $\mathbb{E}[X_n] = 0$ et $\mathbb{E}[X_n^2] = \sigma^2$. Si $S_n = X_1 + \cdots + X_n$, on a par le théorème de la limite centrale

$$\frac{S_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$$

ce qui entraine que, pour toute fonction h continue bornée,

$$\lim_{n\to\infty} \mathbb{E}\Big[h\Big(\frac{S_n}{\sqrt{n}}\Big)\Big] = \int_{\mathbb{R}} h(x)dG(x)$$

où G est la mesure gaussienne $\mathcal{N}(0, \sigma^2)$. On a également, par le théorème de la limite centrale presque sûr (TLCPS), que la mesure emprique aléatoire

$$\frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \delta_{\frac{S_k}{\sqrt{k}}} \Longrightarrow G \quad \text{p.s.}$$

En d'autres termes, pour toute fonction h continue bornée,

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} h\Big(\frac{S_k}{\sqrt{k}}\Big) = \int_{\mathbb{R}} h(x)dG(x) \quad \text{p.s.}$$

Ce théorème a été démontré par [2] et [16, 17], et dans sa forme présente par [11]. Le TLCPS a aussi été établi dans un cadre *martingales* par [3, 4], [5] et [14, 15], plus précisément défini de la manière suivante. Soit (ε_n) une suite de différences de martingales adaptée à une filtration $\mathbb{F} = (\mathcal{F}_n)$. Soit (Φ_n) une suite de variables aléatoires adaptée à $\mathbb{F}$. La transformée de martingale réelle (M_n) est définie, pour tout $n \geq 1$, par

$$M_n \stackrel{\text{def}}{=} \sum_{k=1}^{n} \Phi_{k-1}\varepsilon_k.$$

On définit également le coefficient d'explosion f_n associé à (Φ_n) par

$$f_n \stackrel{\text{def}}{=} \frac{\Phi_n^2}{s_n}, \quad \text{avec} \quad s_n \stackrel{\text{def}}{=} \sum_{k=0}^{n} \Phi_k^2.$$

Le TLCPS pour les martingales [3] peut alors s'exprimer en version simplifiée de la façon suivante. Sous des hypothèses raisonnables de moment et de convergence ad hoc, on a pour toute fonction h continue bornée

$$\lim_{n\to\infty}\frac{1}{\log s_n}\sum_{k=1}^{n} f_k h\Big(\frac{M_k}{\sqrt{s_{k-1}}}\Big) = \int_{\mathbb{R}} h(x)dG(x) \quad \text{p.s.} \tag{1}$$

Une démarche naturelle est de chercher à savoir si cette convergence reste vraie pour des fonctions h non bornées. On trouve dans [1] que, si (ε_n) admet un moment conditionnel d'ordre $> 2p$ fini, alors la convergence (1) est vraie pour $h(x) = x^{2p}$ avec $p \geq 1$.

Théorème 1.1 (Convergence des moments dans le TLCPS scalaire). *On suppose que (ε_n) vérifie $\mathbb{E}[\varepsilon_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2$ p.s. et qu'il existe un entier $p \geq 1$ et un réel $a > 2p$ tels que*

$$\sup_{n\geq 0}\mathbb{E}[|\varepsilon_{n+1}|^a|\mathcal{F}_n] < \infty \quad p.s.$$

Si (f_n) tend vers zéro p.s., on a

$$\lim_{n\to\infty}\frac{1}{\log s_n}\sum_{k=1}^{n} f_k\Big(\frac{M_k^2}{s_{k-1}}\Big)^p = \frac{\sigma^{2p}(2p)!}{2^p p!} \quad p.s. \tag{2}$$

On reconnaît, dans la limite (2), le moment d'ordre $2p$ de la loi gaussienne $\mathcal{N}(0,\sigma^2)$. Dans ce rapport, on propose un résultat analogue au théorème 1.1 de [1] dans un cadre vectoriel.

Soit (M_n) une martingale à valeurs dans $\mathbb{R}^d$ adaptée à une filtration $\mathbb{F}$. On suppose que (M_n) est de carré intégrable. Son processus croissant est la suite $(\langle M\rangle_n)$ de matrices carrées d'ordre d, symétriques et définies positives, données par

$$\langle M\rangle_n = \sum_{k=1}^{n}\mathbb{E}[(M_k - M_{k-1})(M_k - M_{k-1})^t|\mathcal{F}_{k-1}].$$

Un premier TLCPS pour les martingales vectorielles discrètes a été établi dans [5] et [6] sous des conditions assez restrictives sur le processus croissant $(\langle M\rangle_n)$. Notre objectif est de montrer que, sous des hypothèses appropriées sur $(\langle M\rangle_n)$, il y a convergence des moments d'ordre pair dans le TLCPS pour (M_n). On va se placer dans le cadre plus générale des transformées de martingales vectorielles (M_n) de la forme

$$M_n \stackrel{\text{def}}{=} M_0 + \sum_{k=1}^{n}\Phi_{k-1}\varepsilon_k$$

où M_0 peut être choisie arbitrairement et (Φ_n) est une suite de vecteurs aléatoires de $\mathbb{R}^d$ adaptée à $\mathbb{F}$. On note également

$$S_n = \sum_{k=0}^{n} \Phi_k \Phi_k^t + S \tag{3}$$

où S est une matrice déterministe, symétrique et définie positive. Il est clair que, si $\mathbb{E}[\varepsilon_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2$ p.s., alors le processus croissant de (M_n) est donné par $\langle M \rangle_n = \sigma^2 S_{n-1}$. On définit le coefficient d'explosion associé à (Φ_n) par

$$f_n = \Phi_n^t S_n^{-1} \Phi_n = \frac{d_n - d_{n-1}}{d_n} \tag{4}$$

avec $d_n = \det(S_n)$.

Le rapport est organisé de la façon suivante. On énonce dans la section 2 les principaux résultats de convergence pour les transformées de martingales vectorielles. Les applications statistiques sur le calcul des erreurs d'estimation et de prédiction dans les modèles de régressions linéaires font l'objet de la section 3. On étudie en particulier les modèles autorégressifs et les processus de branchement avec immigration. Les démonstrations des résultats ne sont pas données dans leur intégralité, mais les idées clés sont indiquées. On se reportera à [7] pour le détail des preuves.

2 Convergence des moments dans le TLCPS pour les martingales vectorielles

Notre objectif est d'établir une généralisation du théorème 1.1 au cadre vectoriel. Sous des hypothèses de moment appropriées sur (ε_n) ainsi que des hypothèses de convergence raisonnables sur $(\langle M \rangle_n)$, on va montrer la convergence des moments normalisés d'ordre pair dans le TLCPS pour (M_n).

Théorème 2.1. *On suppose que (ε_n) vérifie $\mathbb{E}[\varepsilon_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2$ p.s. et qu'il existe un entier $p \geq 1$ et un réel $a > 2p$ tels que*

$$(H_p) \qquad \sup_{n \geq 0} \mathbb{E}\big[|\varepsilon_{n+1}|^a \big| \mathcal{F}_n\big] < \infty \qquad p.s.$$

On suppose également que (f_n) tend vers zéro p.s. et qu'il existe une suite aléatoire positive (α_n) croissante vers l'infini et une matrice L symétrique et inversible telles que

$$\lim_{n \to \infty} \alpha_n^{-1} S_n = L \quad p.s. \tag{5}$$

Alors, on a presque sûrement

$$\lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} f_k \big(M_k^t S_{k-1}^{-1} M_k\big)^p = \ell(p) = d\sigma^{2p} \prod_{j=1}^{p-1} \big(d+2j\big), \tag{6}$$

$$\lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} \big(M_k^t S_{k-1}^{-1} M_k\big)^p - \big(M_k^t S_k^{-1} M_k\big)^p = \lambda(p) = \frac{p}{d}\ell(p). \tag{7}$$

Remarque 1. La limite $\ell(p)$ correspond au moment d'ordre $2p$ de la norme d'un vecteur gaussien $\mathcal{N}(0, \sigma^2 I_d)$. Par conséquent, le théorème 2.1 établit bien la convergence des moments d'ordre $2p$ dans le TLCPS pour les martingales vectorielles. La normalisation déterministe du TLCPS de [5] est remplacée par la normalisation aléatoire naturelle donnée par le processus croissant.

Remarque 2. L'hypothèse de convergence (5) entraîne que (f_n) tend vers zéro p.s. si et seulement si $\alpha_n \sim \alpha_{n-1}$ p.s. car $\det L > 0$ et

$$\lim_{n\to\infty} \frac{d_n}{\alpha_n^d} = \det L \quad \text{p.s.}$$

Démonstration. La démonstration complète du théorème 2.1 figure dans [7]. On indique seulement ici les idées principales de la preuve. Pour alléger les notations, on commence par définir les variables

$$\begin{aligned} V_n &\stackrel{\text{def}}{=} M_n^t S_{n-1}^{-1} M_n, \qquad (8)\\ \varphi_n &\stackrel{\text{def}}{=} \alpha_n^{-1} \Phi_n^t L^{-1} \Phi_n, \\ v_n &\stackrel{\text{def}}{=} \alpha_{n-1}^{-1} M_n^t L^{-1} M_n. \end{aligned}$$

Tout d'abord, on peut remarquer, en utilisant le symétrie de L, que la convergence (5) entraîne que presque sûrement

$$f_n = \varphi_n + o\big(\varphi_n\big), \tag{9}$$

$$V_n = v_n + o\big(v_n\big). \tag{10}$$

On déduit directement de (10) que $V_n^p = v_n^p + o\big(v_n^p\big)$. Pour déterminer la limite (6), il suffit donc, d'après le lemme de Toeplitz, d'étudier la convergence presque sûre

$$\lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} f_k V_k^p = \lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} \varphi_k v_k^p. \tag{11}$$

On prouve le théorème 2.1 par récurrence sur $p \geq 1$. L'idée de la preuve est inspirée de celle de [1] dans le cas scalaire. Il s'agit dans un premier temps d'écrire une relation de récurrence sur $M_n^t L^{-1} M_n$. Par définition de la suite (M_n), on a la décomposition

$$M_{n+1}^t L^{-1} M_{n+1} = M_n^t L^{-1} M_n + 2\varepsilon_{n+1} \Phi_n^t L^{-1} M_n + \varepsilon_{n+1}^2 \Phi_n^t L^{-1} \Phi_n.$$

On pose

$$\beta_n = \operatorname{tr}\big(L^{-1/2} S_n L^{-1/2}\big), \tag{12}$$

$$\gamma_n = \frac{\beta_n - \beta_{n-1}}{\beta_n} \qquad \text{et} \qquad \delta_n = \frac{M_n^t L^{-1} \Phi_n}{\beta_n}.$$

Si $m_n = \beta_{n-1}^{-1} M_n^t L^{-1} M_n$, on déduit alors de la décomposition ci-dessus que

$$m_{n+1} = (1 - \gamma_n) m_n + 2\delta_n \varepsilon_{n+1} + \gamma_n \varepsilon_{n+1}^2. \tag{13}$$

La preuve du théorème repose ensuite sur le lemme suivant.

Lemme 2.2. *Sous les hypothèses du théorème 2.1, si $g_n = M_n^t S_{n-1}^{-1} \Phi_n$, on a*

$$\lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} \gamma_k m_k^p = \frac{\ell(p)}{d^{p+1}} \quad p.s. \tag{14}$$

$$\lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} (1 - f_k) g_k^2 m_k^{p-1} = \frac{\lambda(p)}{p d^{p-1}} \quad p.s. \tag{15}$$

Si l'on élève l'égalité (13) à la puissance p, on obtient

$$m_{n+1}^p = \sum_{k=0}^{p} \sum_{\ell=0}^{k} 2^{k-\ell} C_p^k C_k^\ell \, \gamma_n^\ell \delta_n^{k-\ell} \big((1-\gamma_n) m_n\big)^{p-k} \varepsilon_{n+1}^{k+\ell}. \tag{16}$$

Après quelques simplifications élémentaires, on trouve la relation de récurrence

$$m_{n+1}^p + \mathcal{A}_n(p) = m_1^p + \mathcal{B}_{n+1}(p) + \mathcal{W}_{n+1}(p) \tag{17}$$

avec

$$\mathcal{A}_n(p) = \sum_{k=1}^{n} \beta_k^{-p} \big(\beta_k^p - \beta_{k-1}^p\big) m_k^p, \quad \mathcal{B}_{n+1}(p) = \sum_{\ell=1}^{2p-1} \sum_{k=1}^{n} b_k(\ell) \varepsilon_{k+1}^\ell,$$

$$\mathcal{W}_{n+1}(p) = \sum_{k=1}^{n} \gamma_k^p \varepsilon_{k+1}^{2p}.$$

Si $1 \leq \ell \leq p-1$, on a

$$b_k(\ell) = \sum_{j=0}^{\lfloor \ell/2 \rfloor} 2^{\ell-2j} C_p^{\ell-j} C_{\ell-j}^j \gamma_k^j \delta_k^{\ell-2j} \big((1-\gamma_k)m_k\big)^{p-\ell+j},$$

tandis que si $p \leq \ell \leq 2p-1$,

$$\begin{aligned} b_k(\ell) = \sum_{j=\ell-(p-1)}^{\lfloor \ell/2 \rfloor} & 2^{\ell-2j} C_p^{\ell-j} C_{\ell-j}^j \gamma_k^j \delta_k^{\ell-2j} \big((1-\gamma_k)m_k\big)^{p-\ell+j} \\ & + C_p^{\ell-p} 2^{2p-\ell} - \delta_k^{2p-\ell} \gamma_k^{\ell-p}. \end{aligned}$$

La preuve consiste en l'étude des comportements asymptotiques de $\mathcal{W}_{n+1}(p)$, $\mathcal{B}_{n+1}(p)$ et m_n^p afin d'en déduire des informations sur $\mathcal{A}_n(p)$. Tout d'abord, pour $p=1$, on a par le lemme de Chow [8, page 22]

$$\lim_{n\to\infty} \frac{1}{\log \beta_n} \mathcal{W}_{n+1}(1) = \sigma^2 \quad \text{p.s.}$$

Il est aussi facile de voir que $\mathcal{B}_{n+1}(1) = o\big(\mathcal{A}_n(1)\big)$ p.s. On déduit également de la relation (2.30) de [19] que $m_{n+1} = o(\log \beta_n)$ p.s. Par suite, on a par l'égalité (17)

$$\lim_{n\to\infty} \frac{1}{\log \beta_n} \mathcal{A}_n(1) = \sigma^2 \quad \text{p.s.}$$

Cependant, il résulte immédiatement de l'hypothèse de convergence (5) que

$$\lim_{n\to\infty} \frac{\beta_n}{\gamma_n} = d \quad \text{p.s.}$$

ce qui implique $\log d_n \sim d \log \beta_n$ p.s. On tire alors

$$\lim_{n\to\infty} \frac{1}{\log d_n} \mathcal{A}_n(1) = \frac{\sigma^2}{d} \quad \text{p.s.}$$

ce qui entraine la convergence (14) pour $p=1$. On suppose ensuite que $p \geq 2$. On trouve par le lemme de Chow que

$$\mathcal{W}_{n+1}(p) = o\big(\log d_n\big) \quad \text{p.s.}$$

On peut également montrer que

$$\mathcal{B}_{n+1}(p) = \frac{p}{d^{p+1}}\ell(p)\log d_n + o\big(\log d_n\big) + o\big(\mathcal{A}_n(p)\big) \quad \text{p.s.}$$

La relation (2.30) de [19] entraine aussi que $m_n^p = o(\log d_n)$. On obtient ainsi

$$\lim_{n\to\infty} \frac{1}{\log d_n}\mathcal{A}_n(p) = \frac{p}{d^{p+1}}\ell(p) \quad \text{p.s.}$$

ce qui achève la preuve de (14) car

$$\lim_{n\to\infty} \frac{1}{\log d_n}\sum_{k=1}^{n}\gamma_k m_k^p = \lim_{n\to\infty} \frac{1}{p\log d_n}\mathcal{A}_n(p) = \frac{\ell(p)}{d^{p+1}} \quad \text{p.s.}$$

Pour la seconde partie du lemme 2.2, on raisonne de la même manière via l'égalité

$$V_{n+1}^p = \sum_{k=0}^{p}\sum_{\ell=0}^{k} 2^{k-\ell}C_p^k C_k^\ell \; f_n^\ell g_n^{k-\ell} h_n^{p-k}\varepsilon_{n+1}^{k+\ell} \tag{18}$$

avec $g_n = \Phi_n^t S_{n-1}^{-1} M_n$ et $h_n = M_n^t S_n^{-1} M_n$. □

Dans la plupart des applications statistiques, l'hypothèse de convergence (5) est vérifiée. Cependant, cette hypothèse technique a été introduite de façon à pouvoir contourner le problème vectoriel en se ramenant d'une certaine façon au cadre scalaire. En effet, l'hypothèse (5) entraine que toutes les valeurs propres de la matrice S_n tendent vers l'infini à la même vitesse α_n. Il est naturel d'espérer pouvoir établir le résultat suivant sans l'hypothèse de convergence (5).

Conjecture 2.3. *On suppose que (ε_n) vérifie (H_p) pour un entier $p \geq 1$. Alors, on a*

$$\sum_{k=1}^{n} f_k\big(M_k^t S_{k-1}^{-1} M_k\big)^p = \mathcal{O}\big(\log d_n\big) \quad \textit{p.s.}$$

$$\sum_{k=1}^{n}\big(M_k^t S_{k-1}^{-1} M_k\big)^p - \big(M_k^t S_k^{-1} M_k\big)^p = \mathcal{O}\big(\log d_n\big) \quad \textit{p.s.}$$

3 Applications

3.1 Modèles de regression linéaire

À partir de nouvelles propriétés asymptotiques presque sûres pour les puissances de transformées de martingales vectorielles, on va établir des résultats de convergence sur les erreurs d'estimation et de prédiction associées aux modèles de régression linéaire. Ces derniers sont définis, pour tout $n \geq 1$, par la relation

$$X_{n+1} = \theta^t \Phi_n + \varepsilon_{n+1} \tag{19}$$

où $\theta \in \mathbb{R}^d$ est le paramètre inconnu du modèle. Les variables aléatoires X_n, Φ_n et ε_n sont respectivement l'observation, le vecteur de régression et le bruit du système. On propose ci-après deux illustrations des résultats de la section 2 portant sur les processus autorégressifs linéaires et ainsi que sur les processus de branchement avec immigration.

Pour une suite d'estimateurs $(\widehat{\theta}_n)$ de θ, on s'intéresse à la performance asymptotique de $\widehat{\theta}_n^t \Phi_n$ comme prédicteur de X_{n+1}. Plus précisément, on se concentre sur l'erreur de prédiction $X_{n+1} - \widehat{\theta}_n^t \Phi_n$ et sur l'erreur d'estimation $\widehat{\theta}_n - \theta$. Il est plus approprié (voir par exemple [10]) de considérer les erreurs cumulées de prédiction et d'estimation, respectivement définies, pour tout $p \geq 1$, par

$$C_n(p) = \sum_{k=0}^{n-1} (X_{k+1} - \widehat{\theta}_k^t \Phi_k)^{2p} \tag{20}$$

et

$$G_n(p) = \sum_{k=1}^{n} k^{p-1} \|\widehat{\theta}_k - \theta\|^{2p}. \tag{21}$$

Dans le cas scalaire $d = 1$, sous des hypothèses de moment appropriées, des résultats asymptotiques sur $(C_n(p))$ et $(G_n(p))$ se trouvent dans [1] avec l'estimateur des moindre carrés

$$\widehat{\theta}_n = S_{n-1}^{-1} \sum_{k=1}^{n} \Phi_{k-1} X_k. \tag{22}$$

Notre but est d'étendre ces résultats au cadre multidimensionnel. Dans ce contexte, il y a dans [8,19] des résultats asymptotiques pour $(C_n(p))$ et $(G_n(p))$ mais seulement dans le cas $p = 1$. Pour la consistance forte de l'estimateur des moindres carrés, on peut également consulter les travaux de [12, 18]. Ces résultats sont étendus dans [9]. D'autre part, le comportement asymptotique de l'estimateur empirique de la covariance associée au modèle (19) est étudié dans [9, 12, 13, 18]. On peut se poser la question générale suivante : comment utiliser les résultats de convergence de la section 2 afin obtenir le comportement asymptotique des suites $(C_n(p))$ et $(G_n(p))$? On peut déduire de (19) et de (22) que

$$\widehat{\theta}_n - \theta = S_{n-1}^{-1} M_n \qquad \text{avec} \qquad M_n = M_0 + \sum_{k=1}^{n} \Phi_{k-1}\varepsilon_k \tag{23}$$

où $M_0 = -S\theta$. On pose alors

$$\pi_n = \big(\theta - \widehat{\theta}_n\big)^t \Phi_n = X_{n+1} - \widehat{\theta}_n^t \Phi_n - \varepsilon_{n+1}. \tag{24}$$

Les relations (23) et (24) permettent d'écrire

$$\pi_n^2 = M_n^t S_{n-1}^{-1} \Phi_n \Phi_n^t S_{n-1}^{-1} M_n.$$

Par ailleurs, en appliquant la formule de Riccati [8], on a

$$S_{n-1}^{-1} = S_n^{-1} + (1 - f_n) S_{n-1}^{-1} \Phi_n \Phi_n^t S_{n-1}^{-1}.$$

Par conséquent,

$$a_n(1) \stackrel{\text{def}}{=} M_n^t S_{n-1}^{-1} M_n - M_n^t S_n^{-1} M_n = (1 - f_n)\pi_n^2. \tag{25}$$

Il est souvent difficile d'obtenir des résultats asymptotiques sur le coefficient d'explosion f_n. Dans les modèles considérés dans ce rapport, (f_n) converge p.s. vers zéro. Pour étudier le comportement asymptotique de $(G_n(p))$ et $(C_n(p))$, on va utiliser des propriétés asymptotiques de $\big(a_n(1)\big)^p$, sous des hypothèses de moments appropriées. Du même coup, on estime les moments d'ordre $2p$ du bruit (ε_n).

Corollaire 3.1. *Sous les hypothèses du théorème 2.1, on a*

$$\lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} \big(a_k(1)\big)^p = \begin{cases} 0 & si \quad p > 1, \\ \sigma^2 & si \quad p = 1. \end{cases} \tag{26}$$

Démonstration. Dans le cas particulier $p = 1$, la convergence (26) correspond exactement à (7). On suppose maintenant $p > 1$. On déduit de l'inégalité

$$a_n(1) \leq f_n V_n$$

et de la convergence presque sûre de (f_n) vers zéro, en appliquant le lemme de Kronecker et le lemme 2.2, que

$$0 \leq \lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} \big(a_k(1)\big)^p \leq \lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} \big(f_k V_k\big)^p = 0 \quad \text{p.s.}$$

□

3.2 Estimation des moments, erreurs d'estimation et de prédiction

On suppose que le bruit (ε_n) vérifie $\mathbb{E}[\varepsilon_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2$ ainsi que (H_1). Si

$$\Delta_n = \frac{1}{n} \sum_{k=1}^{n} \varepsilon_k^2,$$

la loi forte des grands nombres pour les martingales entraine que (Δ_n) converge p.s. vers σ^2. Sous les hypothèses du théorème 2.1 avec $p = 1$, on déduit alors de (26) que l'estimateur de σ^2

$$\Gamma_n = \frac{1}{n} \sum_{k=0}^{n-1} \big(X_{k+1} - \widehat{\theta}_k^t \Phi_k\big)^2$$

est fortement consistant avec

$$\lim_{n\to\infty} \frac{n}{\log d_n} (\Gamma_n - \Delta_n) = \sigma^2 \quad \text{p.s.} \tag{27}$$

On peut maintenant proposer des estimateurs consistants des moments d'ordre supérieur de (ε_n). Pour tout $p \geq 1$,

$$\Gamma_n(2p) = \frac{1}{n} \sum_{k=0}^{n-1} (X_{k+1} - \widehat{\theta}_k^t \Phi_k)^{2p} \tag{28}$$

est un estimateur naturel du moment $\sigma(2p)$ d'ordre $2p$ du bruit (ε_n). On peut remarquer que $n\Gamma_n(2p) = C_n(p)$. Le corollaire suivant donne des résultats asymptotiques pour $\Gamma_n(2p)$.

Corollaire 3.2. *On se place sous les hypothèses du théorème 2.1 avec* $p \geq 2$. *Pour tout entier* $1 \leq q \leq p$ *tel que, pour tout* $n \geq 0$, $\mathbb{E}\big[\varepsilon_{n+1}^{2q}\big|\mathcal{F}_n\big] = \sigma(2q)$, $\Gamma_n(2q)$ *est un estimateur fortement consistant de* $\sigma(2q)$ *et*

$$\Big(\Gamma_n(2q) - \frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^{2q}\Big)^2 = \mathcal{O}\Big(\frac{\log d_n}{n}\Big) \quad \textit{p.s.} \tag{29}$$

Démonstration. On a déjà vu par (27) que le corollaire 3.2 est vrai pour $q = 1$. On peut donc supposer que $q \geq 2$. Si l'on développe l'expression de $\Gamma_n(2q)$, on a par (24)

$$n\big(\Gamma_n(2q) - \Delta_n(2q)\big) = \sum_{k=0}^{n-1}\pi_k^{2q} + \sum_{\ell=1}^{2q-1} C_q^{\ell}\sum_{k=0}^{n-1}\pi_k^{2q-\ell}\varepsilon_{k+1}^{\ell}.$$

On déduit de (25) et de la convergence p.s. de (f_n) vers zéro associés au corollaire 3.1 que

$$\sum_{k=0}^{n}\pi_k^{2q} = o(\log d_n) \quad \text{p.s.}$$

Pour tout $\ell \in \{1, \ldots, 2q-1\}$, on a la décomposition

$$\sum_{k=0}^{n-1}\pi_k^{2q-\ell}\varepsilon_{k+1}^{\ell} = P_n(\ell) + Q_n(\ell),$$

avec

$$P_n(\ell) = \sum_{k=0}^{n-1}\pi_k^{2q-\ell}\sigma_k(\ell) \qquad \text{et} \qquad Q_n(\ell) = \sum_{k=0}^{n-1}\pi_k^{2q-\ell}e_{k+1}(\ell),$$

où $\sigma_n(\ell) = \mathbb{E}\big[\varepsilon_{n+1}^{\ell}\big|\mathcal{F}_n\big]$ et $e_{n+1}(\ell) = \varepsilon_{n+1}^{\ell} - \sigma_n(\ell)$. Tout d'abord, puisque les moments $\sigma_n(\ell)$ sont presque sûrement bornés, on a

$$|P_n(\ell)| = \mathcal{O}\Big(\sum_{k=0}^{n-1}\pi_k^{2q-\ell}\Big) = \mathcal{O}(\log d_n) \quad \text{p.s.}$$

On peut également montrer qu'on a l'estimation

$$|Q_n(\ell)|^2 = \mathcal{O}(n\log d_n) \quad \text{p.s.}$$

Par conséquent, on obtient

$$n^2\big(\Gamma_n(2q) - \Delta_n(2q)\big)^2 = \mathcal{O}\big(n\log d_n\big) \quad \text{p.s.},$$

ce qui achève la preuve du corollaire 3.2. □

Il est maintenant possible de déduire du corollaire 3.2 le comportement asymptotique de $(C_n(q))$. Sous les hypothèses du corollaire 3.2, la convergence (29) implique que $C_n(q)/n$ converge p.s. vers $\sigma(2q)$. De plus, comme (ε_n) admet un moment conditionnel fini d'ordre $a > 2p$, on peut déduire du lemme de Chow que, pour tout c tel que $2pa^{-1} < c < 1$,

$$\left| \frac{1}{n} \sum_{k=1}^{n} \varepsilon_k^{2q} - \sigma(2q) \right| = o(n^{c-1}) \quad \text{p.s.} \tag{30}$$

Il découle alors de (29) et (30) que, si $\log d_n = o\big(n^c\big)$,

$$\left| \frac{1}{n}\, C_n(q) - \sigma(2q) \right|^2 = o(n^{c-1}) \quad \text{p.s.}$$

Avant d'obtenir des résultats sur les erreurs cumulées d'estimation $(G_n(p))$, on énonce un autre corollaire du théorème 2.1.

Corollaire 3.3. *Sous les hypothèses du théorème 2.1, on a*

$$\lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} f_k \left((\widehat{\theta}_k - \theta)^t S_k (\widehat{\theta}_k - \theta) \right)^p = \ell(p) \quad \textit{p.s.} \tag{31}$$

De plus, on suppose qu'il existe une matrice inversible L telle que

$$\lim_{n\to+\infty} \frac{1}{n}\, S_n = L \qquad \textit{p.s.} \tag{32}$$

Alors, on a

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} k^{p-1} \left((\widehat{\theta}_k - \theta)^t L (\widehat{\theta}_k - \theta) \right)^p = \ell(p) \quad \textit{p.s.} \tag{33}$$

Remarque 3. Comme L est définie positive, on tire immédiatement de (33) que

$$G_n(p) = \mathcal{O}\big(\log n\big) \quad \text{p.s.}$$

Démonstration. On montre facilement à partir de la définition de S_n et de l'estimateur des moindres carrés $\widehat{\theta}_n$ que

$$\big(\widehat{\theta}_n - \theta\big)^t S_n \big(\widehat{\theta}_n - \theta\big) = V_n + g_n^2.$$

Ainsi, on déduit de la convergence (6), de la limite nulle du coefficient d'explosion f_n et du lemme de Kronecker que presque sûrement

$$\lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} f_k \Big(\big(\widehat{\theta}_k - \theta\big)^t S_k \big(\widehat{\theta}_k - \theta\big) \Big)^p = \lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} f_k V_k^p = \ell(p).$$

Ainsi la convergence (31) est une conséquence immédiate du théorème 2.1. Pour la deuxième partie du corollaire, on peut déjà montrer en utilisant la racine carrée de L que

$$\Big(\big(\hat{\theta}_n - \theta\big)^t L \big(\hat{\theta}_n - \theta\big) \Big)^p \sim \big(n^{-2} m_n \beta_n\big)^p \sim \big(d^2 m_n \beta_n^{-1}\big)^p \quad \text{p.s.}$$

Ainsi, on obtient

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} k^{p-1} \Big(\big(\hat{\theta}_k - \theta\big)^t L \big(\hat{\theta}_k - \theta\big) \Big)^p = \lim_{n\to\infty} \frac{d^{p+1}}{\log n} \sum_{k=1}^{n} \frac{m_k^p}{\beta_k} \quad \text{p.s.}$$

On cherche à comparer cette somme avec celle étudiée dans le lemme 2.2. Grâce à une transformation d'Abel, on a la décomposition

$$\sum_{k=1}^{n} \gamma_k m_k^p = \frac{m_n^p}{\beta_n} \big(\Sigma_n - d(n-1)\big) - \frac{m_1^p}{\beta_0} \Sigma_0 + r_n + d \sum_{k=1}^{n-1} \frac{m_k^p}{\beta_k}, \tag{34}$$

avec

$$\Sigma_n \stackrel{\text{def}}{=} \sum_{k=1}^{n} \beta_k \gamma_k = \sum_{k=1}^{n} \Phi_k^t L^{-1} \Phi_k \sim \beta_n$$

et

$$r_n \stackrel{\text{def}}{=} \sum_{k=1}^{n-1} \Big(\frac{m_k^p}{\beta_k} - \frac{m_{k+1}^p}{\beta_{k+1}} \Big) \big(\Sigma_k - kd\big).$$

On peut voir aisément que

$$\frac{m_n^p}{\beta_n} \big(\Sigma_n - d(n-1)\big) - \frac{m_1^p}{\beta_0} \Sigma_0 = o\big(\log d_n\big) \quad \text{p.s.}$$

Il reste à prouver que $r_n = o\big(\log n\big)$ p.s. En effet, on aura alors par le lemme 2.2 que

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{m_k^p}{\beta_k} = \frac{1}{d} \lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \gamma_k m_k^p = \frac{\ell(p)}{d^{p+1}} \quad \text{p.s.}$$

On écrit r_n comme somme deux termes :

$$r_n = \sum_{k=1}^{n-1} \frac{\Sigma_k - kd}{\beta_k}\big(m_k^p - m_{k+1}^p\big) + \sum_{k=1}^{n-1} \frac{\Sigma_k - kd}{\beta_k} \gamma_{k+1} m_{k+1}^p. \tag{35}$$

Ensuite, en utilisant la preuve du théorème 2.1 et la décomposition (16), on trouve pour le premier terme

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n-1} \frac{\Sigma_k - kd}{\beta_k}\Big(\beta_k^{-p}\big(\beta_k^p - \beta_{k-1}^p\big)m_k^p - w_{k+1} - b_{k+1}\Big) = 0 \quad \text{p.s.}$$

On déduit directement du lemme 2.2 et de la convergence presque sûre de $\Sigma_n - nd/\beta_n$ vers zéro, que le deuxième terme est aussi un $o(\log n)$ p.s. Finalement, on obtient

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} k^{p-1}\Big(\big(\hat{\theta}_k - \theta\big)^t L\big(\hat{\theta}_k - \theta\big)\Big)^p = \ell(p) \quad \text{p.s.}$$

□

On va maintenant appliquer ces résultats aux processus autorégressifs linéaires ainsi qu'aux processus de branchement avec immigration, qui sont des cas particuliers du modèle de régression (19).

3.3 Processus autorégressif linéaire

Le processus autorégressif linéaire est un cas particulier du modèle de régression (19). Il est défini pour tout $n \geq 1$, par

$$X_{n+1} = \sum_{k=1}^{d} \theta_k X_{n-k+1} + \varepsilon_{n+1}. \tag{36}$$

La matrice compagne C associée à ce processus est donnée par

$$C = \begin{pmatrix} \theta_1 & \theta_2 & \dots & \theta_{d-1} & \theta_d \\ 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

On s'intéresse ici seulement au *cas stable*, c'est-à-dire quand $\rho(C) < 1$, où $\rho(C)$ désigne le rayon spectral de la matrice compagne C. On suppose que (ε_n) est soit une suite de variables aléatoires indépendantes et de même loi centrée et de variance σ^2, soit une suite de différences de martingales satisfaisant $\mathbb{E}[\varepsilon_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2$ avec un moment conditionnel d'ordre > 2 fini. Si

$$\Gamma = \sigma^2 \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$$

le résultat de convergence (32) est vérifié avec L donnée par

$$L = \sum_{k=0}^{\infty} C^k \Gamma (C^t)^k. \tag{37}$$

De plus, on peut facilement voir que L est inversible. Par conséquent, $\log d_n$ est p.s. équivalent à $d \log n$. Sous les hypothèses du corollaire 3.2, $\Gamma_n(2q)$ est un estimateur fortement consistant de $\sigma(2q)$,

$$\Big(\Gamma_n(2q) - \frac{1}{n}\sum_{k=1}^{n} \varepsilon_k^{2q}\Big)^2 = \mathcal{O}\Big(\frac{\log n}{n}\Big) \quad \text{p.s.} \tag{38}$$

et le corollaire 3.3 s'applique en remplaçant $\log d_n$ par $d \log n$. Sous les hypothèses du corollaire 3.3, on a aussi

$$G_n(p) = \mathcal{O}(\log n) \quad \text{p.s.}$$

Finalement, la convergence (33) est vérifiée pour L donnée par (37).

3.4 Processus de branchement avec immigration

3.4.1 Estimation de la moyenne

On considère un processus de branchement (X_n) sujet à une composante d'immigration indépendante à chaque génération. La notion d'immigration correspond au fait que la population de référence peut s'enrichir d'apports extérieurs : par exemple, on peut modéliser l'évolution d'un patrimoine génétique, de phénomènes en écologie, en physique des particules ou en épidémiologie. Le processus de branchement (X_n) est donné par la relation de récurrence

$$X_{n+1} = \sum_{k=1}^{X_n} Y_{n,k} + I_{n+1} \tag{39}$$

avec $X_0 = 1$. La variable aléatoire I_n correspond à l'effectif de l'immigration à la n^e génération tandis que, pour tout $1 \leq k \leq X_n$, $Y_{n,k}$ désigne le nombre de descendants du k^e individu de la n^e génération. On suppose que les suites $(Y_{n,k})$ et (I_n) sont indépendantes et constituées de variables aléatoires indépendantes et de même loi à valeurs dans $\mathbb{N}$. On pose

$$\begin{aligned} \mathbb{E}[Y_{n,k}] &= m, & \mathbb{E}[I_n] &= \lambda, \\ \text{var}[Y_{n,k}] &= \sigma^2, & \text{var}[I_n] &= b^2. \end{aligned}$$

La relation de récurrence (39) peut s'écrire sous la forme autorégressive

$$X_{n+1} = mX_n + \lambda + \varepsilon_{n+1}, \tag{40}$$

où (ε_n) est une suite de différences de martingales. Le processus (X_n) est donc un cas particulier de (19) avec $\Phi_n^t = (X_n, 1)$ et $\theta^t = (m, \lambda)$. Cependant, dans ce modèle, le moment conditionnel d'ordre 2 du bruit n'est pas presque sûrement borné car

$$\mathbb{E}\big[\varepsilon_{n+1}^2 \mid \mathcal{F}_n\big] = \sigma^2 X_n + b^2.$$

Pour cette raison, on considère le modèle de régression suivant

$$\widetilde{X}_{n+1} = \theta^t \widetilde{\Phi}_n + \widetilde{\varepsilon}_{n+1},$$

où les variables aléatoires $\widetilde{X}_{n+1}$, Φ_n et $\widetilde{\varepsilon}_n$ sont données par

$$\begin{aligned} \widetilde{X}_{n+1} &= c_n^{-1/2} X_{n+1}, & \widetilde{\Phi}_n &= c_n^{-1/2} \Phi_n, \\ \widetilde{\varepsilon}_{n+1} &= c_n^{-1/2} \varepsilon_{n+1}, & c_n &= X_n + 1. \end{aligned}$$

Avec ce modèle, on a clairement la majoration

$$\mathbb{E}\big[\widetilde{\varepsilon}_{n+1}^2 \mid \mathcal{F}_n\big] \leq \sigma^2 + b^2.$$

De plus, dans le cas stable $m < 1$, la convergence suivante a été établie dans [20]

$$\lim_{n\to\infty} n^{-1} \widetilde{S}_n \stackrel{\text{def}}{=} \lim_{n\to\infty} n^{-1} \big(S + \sum_{k=0}^{n} c_k^{-1} \Phi_k \Phi_k^t\big) = L$$

où L est la matrice définie positive donnée par

$$L = \begin{pmatrix} \mathbb{E}\Big[\frac{X^2}{X+1}\Big] & \mathbb{E}\Big[\frac{X}{X+1}\Big] \\ \\ \mathbb{E}\Big[\frac{X}{X+1}\Big] & \mathbb{E}\Big[\frac{1}{X+1}\Big] \end{pmatrix}.$$

La notation X désigne une variable aléatoire suivant la loi stationnaire associée à la chaîne de Markov (X_n). Par conséquent, les hypothèses des corollaires 3.2 et 3.3 sont vérifiées. On obtient donc

$$\Big(\Gamma_n(2q) - \frac{1}{n}\sum_{k=1}^{n} \varepsilon_k^{2q}\Big)^2 = \mathcal{O}\Big(\frac{\log n}{n}\Big) \quad \text{p.s.}$$

$$\left|\frac{1}{n}\ C_n(q) - \sigma(2q)\right|^2 = o(n^{c-1}) \quad \text{p.s.}$$

$$G_n(p) = \mathcal{O}(\log n) \quad \text{p.s}$$

3.4.2 Estimation de la variance

La définition de (ε_n) dans la relation (40) permet d'écrire la décomposition

$$\varepsilon_{n+1}^2 = \sigma^2 X_n + b^2 + v_{n+1},$$

où (v_n) est une suite de différences de martingales vérifiant

$$\mathbb{E}\big[v_{n+1}^2|\mathcal{F}_n\big] = 2\sigma^4 X_n^2 + X_n(\tau^4 - 3\sigma^4 + 4b^2\sigma^2) + \nu^4 - b^4,$$

en notant τ^4 et ν^4 les moments centrés d'ordre 4 de $(Y_{n,k})$ et (I_n). Par conséquent, on a la majoration

$$\mathbb{E}\big[c_n^{-2} v_{n+1}^2|\mathcal{F}_n\big] \leq \tau^4 + 4b^2\sigma^2 + \nu^4.$$

Comme la moyenne θ est inconnue, on utilise l'estimateur des variances $\eta^t = (\sigma^2, b^2)$ défini par

$$\widehat{\eta}_n = Q_n^{-1} \sum_{k=1}^{n} c_k^{-2} \Phi_k \widehat{\varepsilon}_{k+1}^2, \quad \widehat{\varepsilon}_{k+1} = X_{k+1} - \widehat{\theta}_k \Phi_k,$$

où

$$Q_n = S + \sum_{k=1}^{n} c_k^{-2} \Phi_k \Phi_k^t.$$

D'après [20], dans le cas stable $m < 1$, l'hypothèse de convergence (32) est vérifiée puisque

$$\lim_{n\to\infty} n^{-1}Q_n = \Lambda$$

où Λ est la matrice définie positive donnée par

$$\Lambda = \begin{pmatrix} \mathbb{E}\Big[\frac{X^2}{(X+1)^2}\Big] & \mathbb{E}\Big[\frac{X}{(X+1)^2}\Big] \\ \\ \mathbb{E}\Big[\frac{X}{(X+1)^2}\Big] & \mathbb{E}\Big[\frac{1}{(X+1)^2}\Big] \end{pmatrix}.$$

Ainsi les hypothèses des corollaires 3.2 et 3.3 sont également vérifiées. Les comportements asymptotiques des erreurs d'estimation et de prédiction cumulées sont donc identiques au cas de l'estimation de la moyenne.

Références

[1] Bercu, B. On the convergence of moments in the almost sure central limit theorem for martingales with statistical applications. *Stochastic Processes and their applications 111* (2004), 157–173.

[2] Brosamler, G. A. An almost everywhere central limit theorem. *Math. Proc. Cambridge Philos. Soc. 104*, 3 (1988), 561–574.

[3] Chaâbane, F. Version forte du théorème de la limite centrale fonctionnel pour les martingales. *C. R. Acad. Sci. Paris Sér. I Math. 323*, 2 (1996), 195–198.

[4] Chaâbane, F. Invariance principles with logarithmic averaging for martingales. *Studia Sci. Math. Hungar. 37*, 1-2 (2001), 21–52.

[5] Chaâbane, F., and Maâouia, F. Théorèmes limites avec poids pour les martingales vectorielles. *ESAIM Probab. Statist. 4* (2000), 137–189 (electronic).

[6] Chaâbane, F., Maâouia, F., and Touati, A. Généralisation du théorème de la limite centrale presque-sûr pour les martingales vectorielles. *C. R. Acad. Sci. Paris Sér. I Math. 326*, 2 (1998), 229–232.

[7] Cénac, P. *Etude statistique de séquences biologiques et convergence de martingales.* PhD thesis, INRIA Rocquencourt - Université Paul Sabatier Toulouse III, 2006.

[8] Duflo, M. *Random Iterative Methods.* Springer-Verlag, 1997.

[9] Duflo, M., Senoussi, R., and Touati, A. Propriétés asymptotiques presque sûres de l'estimateur des moindres carrés d'un modèle autoregressif vectoriel. *Ann. Inst. Henri Poincaré 27*, 1 (1991), 1–25.

[10] Goodwin, G., and Sin, K. *Adaptative Filtering Prediction and Control.* Prentice-Hall, Englewood Cliffs, N.J., 1984.

[11] Lacey, M. Laws of the iterated logarithm for partial sum processes indexed by functions. *J. Theoret. Probab. 2*, 3 (1989), 377–398.

[12] Lai, T., and Wei, C. Least-squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *The Annals of Statistics 10*, 1 (1982), 154–166.

[13] Lai, T., and Wei, C. Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters. *Journal of Multivariate Analysis 13* (1983), 1–23.

[14] Lifshits, M. A. Lecture notes on almost sure limit theorems. *Publications IRMA 54* (2001), 1–25.

[15] Lifshits, M. A. Almost sure limit theorem for martingales. In *Limit theorems in probability and statistics, Vol. II (Balatonlelle, 1999).* János Bolyai Math. Soc., Budapest, 2002, pp. 367–390.

[16] Schatte, P. On strong versions of the central limit theorem. *Math. Nachr. 137* (1988), 249–256.

[17] Schatte, P. On the central limit theorem with almost sure convergence. *Probab. Math. Statist. 11*, 2 (1990), 237–246 (1991).

[18] Wei, C. Asymptotic properties of least-squares estimates in stochastic regression models. *The Annals of Statistics 13*, 4 (1985), 1498–1508.

[19] Wei, C. Adaptative prediction by least squares predictors in stochastic regression models with applications to time series. *The Annals of Statistics 15*, 4 (1987), 1667–1682.

[20] Wei, C. Z., and Winnicki, J. Estimation of the means in the branching process with immigration. *Ann. Statist. 18*, 4 (1990), 1757–1773.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399