C. R. Acad. Sci. Paris, Ser. I 335 (2002) 267–270

Statistique/*Statistics*
(Probabilités/*Probability Theory*)

Estimation d'une rupture en dépendance faible

Patrick Ango Nze [a], **Clémentine Prieur** [b]

[a] Université Lille 3, UFR AES, BP 149, 59653, Villeneuve d'Ascq cedex, France

[b] Université Cergy-Pontoise, Mathématiques, St Martin, 2, avenue A. Chauvin, 95302 Cergy-Pontoise cedex, France

Reçu le 14 décembre 2001 ; accepté le 31 mai 2002

Note présentée par Paul Deheuvels.

Résumé On se place dans un cadre de dépendance faible proposé par Doukhan et Louhichi dans [2]. On estime le saut au point de rupture, connu, de la fonction de régression du modèle $Y_n = m(X_n) + \sigma(X_n)\varepsilon_n$. La suite $(\varepsilon_n)_{n\in\mathbb{N}}$ est indépendante et identiquement distribuée (i.i.d.) et indépendante de la suite faiblement dépendante $(X_n)_{n\in\mathbb{N}}$. On énonce un théorème de limite centrale du saut. ***Pour citer cet article : P. Ango Nze, C. Prieur, C. R. Acad. Sci. Paris, Ser. I 335 (2002) 267–270.*** © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Change point estimation for a weakly dependent sequence

Abstract Let $(X_n, Y_n)_{n\in\mathbb{N}}$ be a stationary sequence governed by the model $Y_n = m(X_n) + \sigma(X_n)\varepsilon_n$ where $(\varepsilon_n)_{n\in\mathbb{N}}$ is i.i.d. and independent from $(X_n)_{n\in\mathbb{N}}$. The latter sequence satisfy a weak dependence condition proposed by Doukhan and Louhichi in [2]. We provide a Central Limit Theorem for jumps in the regression function. Our method deals with linear local regression described in [4]. We use a variation on Lindeberg–Rio method as in [5]. ***To cite this article: P. Ango Nze, C. Prieur, C. R. Acad. Sci. Paris, Ser. I 335 (2002) 267–270.*** © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

1. Introduction, notations

Dans cette Note, nous nous intéressons au problème d'estimation d'une rupture dans un modèle de régression : $Y_n = m(X_n) + \sigma(X_n)\varepsilon_n$ où la suite $(\varepsilon_n)_{n\in\mathbb{N}}$ est i.i.d. et indépendante de la suite $(X_n)_{n\in\mathbb{N}}$. On reprend l'étude de ce modèle par Grégoire et Hamrouni [4] dans le cadre indépendant, mais dans un contexte de dépendance faible : la suite $(X_n)_{n\in\mathbb{N}}$ est supposée faiblement dépendante au sens de la Définition 1.1. La première partie est consacrée à la présentation du modèle étudié. On suppose que la fonction de régression présente un unique point de discontinuité τ connu. On cherche à estimer l'amplitude du saut. La méthode d'estimation retenue est une méthode non paramétrique. On estime le saut au point τ, $\gamma(\tau) = m_+(\tau) - m_-(\tau)$, en utilisant un estimateur à droite pour $m_+(\tau)$ et un autre à gauche pour $m_-(\tau)$. Les estimateurs à droite et à gauche choisis sont des estimateurs de régression linéaire locale. L'article [4] motive l'utilisation de cette procédure en vue de résoudre des problèmes de rupture. Les résultats de convergence pour l'estimateur de l'amplitude du saut sont présentés dans la deuxième partie. On montre en particulier le Théorème de limite centrale 2.3.

Adresses e-mail : angonze@univ-lille3.fr (P. Ango Nze); prieur@math.u-cergy.fr (C. Prieur).

S1631-073X(02)02460-3/FLA

1.1. Hypothèses de dépendance faible

Commençons par introduire la notion de dépendance faible utilisée dans cette note. Pour un espace mesurable et normé $(E, \|\cdot\|)$ on définit la famille, $\mathbb{L}^\infty(E^u)$ $(u \in \mathbb{N}^*)$, des fonctions numériques mesurables et bornées sur l'ensemble E^u. Ce dernier est muni de la norme l^1 : $\|x\|_1 = \|x_1\| + \cdots + \|x_u\|$, $x = (x_1, \ldots, x_u) \in E^u$.

Le coefficient de Lipschitz d'une fonction $f : E^u \longrightarrow \mathbb{R}$ étant défini par $\mathrm{Lip}(f) = \sup_{x \neq y} \frac{|f(x)-f(y)|}{\|x-y\|_1}$, on s'attache aux fonctions appartenant à la classe $\mathcal{L} = \bigcup_{u=1}^\infty \{f \in \mathbb{L}^\infty(E^u)\text{: } \mathrm{Lip}(f) < \infty,\ \|f\|_\infty \leqslant 1\}$.

La définition de dépendance faible qui suit reprend l'article [2].

DÉFINITION 1.1. – Une suite $(Z_n)_{n\in\mathbb{N}}$ de v.a. à valeurs dans $E = \mathbb{R}^D$ sera dite $(\theta, \mathcal{L}, \psi)$-faiblement dépendante s'il existe une suite $\theta = (\theta_r)_{r\in\mathbb{N}}$ décroissant vers zéro, une fonction numérique ψ définie pour les quadruplets $(f, g, u, v) \in \mathcal{L}^2 \times \mathbb{N}^{*2}$, telles que pour tout triplet $(u, v, r) \in (\mathbb{N}^*)^2 \times \mathbb{N}$, $|\mathrm{Cov}(f(Z_{i_1}, \ldots, Z_{i_u}), g(Z_{j_1}, \ldots, Z_{j_v}))| \leqslant \psi(f, g, u, v)\theta_r$ pour tout u-uplet $(i_1, \ldots, i_u)$ et tout v-uplet $(j_1, \ldots, j_v)$ dès lors que $i_1 \leqslant \cdots \leqslant i_u < i_u + r \leqslant j_1 \leqslant \cdots \leqslant j_v$.

Cette définition présente une alternative intéressante à des conditions de dépendance plus classiques (le mélange, par exemple). Parmi les choix les plus courants de $\psi(f, g, u, v)$, on retient $\psi_1(f, g, u, v) = u\,\mathrm{Lip}(f) + v\,\mathrm{Lip}(g)$, $\psi_1'(f, g, u, v) = v\,\mathrm{Lip}(g)$. La ψ_1 dépendance faible s'applique à des modèles markoviens généraux ainsi qu'aux schémas de Bernoulli. La fonction ψ_1' renvoie à l'analogue causal de ψ_1. L'article [2] propose une revue plus détaillée de tels modèles applicatifs. Le lecteur intéressé pourra le consulter avec profit.

1.2. Hypothèses sur le modèle

Les observations forment une suite $(X_n, Y_n)_{n\in\mathbb{N}}$ de v.a. à valeurs dans $\mathbb{R}^2$. Elles obéissent au modèle

$Y_i = m(X_i) + \sigma(X_i)\varepsilon_i$ pour une fonction $m(\cdot)$ régulière partout sauf au point connu τ, une fonction $\sigma(\cdot)$ continûment dérivable, et une suite de v.a. à valeurs réelles $(\varepsilon_n)_{n\in\mathbb{N}}$ i.i.d. centrée réduite. On suppose de plus que les suites $(X_n)_{n\in\mathbb{N}}$ et $(\varepsilon_n)_{n\in\mathbb{N}}$ sont mutuellement indépendantes.

Les hypothèses suivantes seront utilisées par la suite.

1. $\mathbb{E}\varepsilon_0 = 0$, $\mathbb{E}(\varepsilon_0^2) = 1$, $\mathbb{E}|\varepsilon_0|^3 < \infty$.
2. La loi de X_0 est concentée sur l'intervalle $[0, 1]$. Elle possède une densité f strictement positive et continue.
3. Les lois des couples (X_0, X_k) possèdent des densités $f_k(x, y)$ uniformément bornées par rapport à k : $\sup_{k\geqslant 1} \sup_{x,y} |f_k(x, y)| < \infty$.
4. La fonction de régression $m(\cdot) = \mathbb{E}(Y_0|X_0 = \cdot)$ est deux fois continûment dérivable dans chaque intervalle $[0, \tau]$ et $[\tau, 1]$. En particulier m et ses deux dérivées successives possèdent des limites à gauche et à droite au point τ.
5. La variance conditionnelle $\sigma^2(\cdot) = \mathrm{Var}(Y_0|X_0 = \cdot)$ est continûment dérivable.
6. Le point de saut τ de la fonction $m(\cdot)$ est dans l'intervalle $]\,0, 1[$.

1.3. Estimateurs linéaires locaux

Les fonctions m et σ sont inconnues. On utilise des estimateurs à noyau pour les estimer. Les définitions qui suivent reprennent [3]. Soit K un noyau, et $h_n = h$ une fenêtre. Les paramètres réels $\hat\alpha_x$ et $\hat\beta_x$ sont obtenus en résolvant le problème de minimisation de $\sum_{i=1}^n (Y_i - \alpha_x - \beta_x(x - X_i))^2 K_i(x)$, avec $K_i(x) = K(\frac{x-X_i}{h})$.

DÉFINITION 1.2. – L'estimateur linéaire local de $m(\cdot)$ est $\widehat{m}(x) = \hat\alpha_x = (\sum_{i=1}^n \omega_i(x)Y_i)/\sum_{i=1}^n \omega_i(x)$, avec $\omega_i(x) = K_i(x)(S_2(x) - (x - X_i)S_1(x))$ où, pour $l \in \mathbb{N}$, $S_l(x) = \sum_{i=1}^n (x - X_i)^l K_i(x)$.

1.4. Estimateurs à droite et à gauche

On reprend ici des hypothèses de [4] pour faciliter la lecture du texte. On considère un noyau $K_+ : [-1, 0] \to \mathbb{R}_+$ continûment différentiable. On pose alors

$$K_l^+ = \int_{-1}^{0} x^l K_+(x)\,\mathrm{d}x, \quad \text{et} \quad L_l^+ = \int_{-1}^{0} x^l K_+^2(x)\,\mathrm{d}x, \quad l \in \mathbb{N},$$

$$B_+ = \left(K_2^+\right)^2 - K_3^+ K_1^+, \quad \text{et} \quad V_+ = \int_{-1}^{0} \left(K_2^+ - x K_1^+\right)^2 K_+^2(x)\,\mathrm{d}x.$$

On définit $K_-(x) = K_+(-x)$, puis les quantités K_l^-, L_l^-, B_- et V_- relativement au noyau K_-. On a alors $V_+ = V_-$, $B_+ = B_-$ et $V_+ = (K_2^+)^2 L_0^+ + (K_1^+)^2 L_2^+ - 2K_1^+ K_2^+ L_1^+$.

D'autre part, il n'y a aucune perte de généralité à imposer la normalisation $K_2 K_0 - K_1^2 = 1$. Les noyaux K_+ et K_- conduisent alors à des estimateurs linéaires locaux $\widehat{m}_+$ et $\widehat{m}_-$ définis par la Définition 1.2.

2. Résultats

On suppose vérifiées les hypothèses de la Section 1.2.

THÉORÈME 2.1. – *On suppose que le processus* **X** *est* $(\theta, \mathcal{L}, \psi_1)$*-faiblement dépendant. Si* $\theta_r = r^{-a}$ *avec* $a \geqslant 3$, *si* $h_n \to 0$, *et* $nh_n \to +\infty$, *alors, pour tout* $l \in \mathbb{N}$, *et pour tout* $x \in]0, 1[$

$$\begin{aligned}
S_l(x) &= nh^{l+1} f(x) K_l \bigl(1 + \mathrm{o}_P(1)\bigr),\\
R_l(x) &= \sum_{i=1}^{n} (x - X_i)^l K_i(x) \bigl(m(X_i) - m(x) - (X_i - x) m'(x)\bigr)\\
&= \frac{1}{2} nh^{l+3} m''(x) f(x) K_{l+2} \bigl(1 + \mathrm{o}_P(1)\bigr), \quad x \neq \tau,\\
T_l(x) &= \sum_{i=1}^{n} (x - X_i)^l K_i^2(x) \sigma^2(X_i) = nh^{l+1} \sigma^2(x) f(x) L_l \bigl(1 + \mathrm{o}_P(1)\bigr),\\
\xi_l(x) &= \sum_{i=1}^{n} (x - X_i)^l K_i(x) \sigma(X_i) \varepsilon_i = \mathrm{O}_P\bigl(n^{1/2} h^{l+(1/2)}\bigr),
\end{aligned}$$

$$\sum_{i=1}^{n} \omega_i(x) = n^2 h^4 f^2(x) \bigl(1 + \mathrm{o}_P(1)\bigr). \tag{1}$$

Ligne de démonstration. – On pose $S_l(x) = \sum_{i=1}^{n} Z_i(x)$ et $A_i(x) = \mathrm{Cov}(Z_1(x), Z_{i+1}(x))$. Alors,

$$S_l(x) = n\mathbb{E}(Z_1(x)) + \mathcal{O}_P\Bigl(\sqrt{\mathrm{Var}(S_l(x))}\Bigr). \tag{2}$$

Grâce à l'hypothèse d'existence et de bornitude des densités jointes f_j d'une part, et à l'hypothèse de $(\theta, \mathcal{L}, \psi_1)$ dépendance faible d'autre part, il vient $|\sum_{i=1}^{n-1} (n - i) A_i(x)| \leqslant Cnh^{2l+2-(3/a)}$. On conclue avec (2). L'étude de R_l, T_l et ξ_l est analogue. On déduit ensuite (1) de l'expression de S_l. □

Remarque. – Dans le cas où $h = n^{-\delta}$ ($\delta \in]0, 1[$), la condition sur le taux de dépendance s'énonce $\theta_r = r^{-a}$ avec $a > 3\delta$.

THÉORÈME 2.2. – *On suppose que le processus* **X** *est* $(\theta, \mathcal{L}, \psi_1)$ *faiblement dépendant. Si le paramètre de fenêtre vérifie* $h_n \to 0$ *et* $nh_n \to +\infty$, *si de plus* $\theta_r = r^{-a}$ *avec* $a \geqslant 3$, *alors l'erreur quadratique conditionnelle moyenne de* $\widehat{\gamma}(\tau)$ *est*

$\mathbb{E}[(\widehat{\gamma}(\tau) - \gamma(\tau))^2 | X_1, \ldots, X_n] = [\frac{\gamma^{(2)}(\tau)}{2} h_n^2 B_+]^2 + \frac{2\sigma^2(\tau) V_+}{nh_n f(\tau)} + \mathrm{o}_P(h_n^4 + \frac{1}{nh_n})$.

Démonstration. – La décomposition de l'erreur quadratique conditionnelle moyenne fait ressortir les termes suivants : $\mathbb{E}[(\widehat{m}_+(\tau) - m_+(\tau))^2 | X_1, \ldots, X_n]$, $\mathbb{E}[(\widehat{m}_-(\tau) - m_-(\tau))^2 | X_1, \ldots, X_n]$, et $\mathbb{E}[(\widehat{m}_+(\tau) - m_+(\tau))(\widehat{m}_-(\tau) - m_-(\tau)) | X_1, \ldots, X_n]$.

Les deux premiers termes se traitent comme dans le cadre indépendant. On suit les décompositions de [4] :

$$\mathbb{E}\big[\big(\widehat{m}_+(\tau)-m_+(\tau)\big)^2|X_1,\ldots,X_n\big]=\left(\frac{\sum_{i=1}^n \omega_i^+(\tau)(m_+(X_i)-m_+(\tau))}{\sum_{i=1}^n \omega_i^+(\tau)}\right)^2+\frac{\sum_{i=1}^n(\omega_i^+(\tau))^2\sigma^2(X_i)}{(\sum_{i=1}^n \omega_i^+(\tau))^2}$$
$$=\big(B_n^+\big)^2+V_n^+.$$

On établit une formule analogue avec m_-. Par contre, on n'a plus ici l'indépendance de $\widehat{m}_+(\tau)-m_+(\tau)$ et de $\widehat{m}_-(\tau)-m_-(\tau)$ conditionnellement à $X_1,\ldots,X_n$. Cependant, remarquons que conditionnellement à $X_1,\ldots,X_n$, $\omega_i^+(\tau)\omega_i^-(\tau)=0$ presque sûrement : c'est une conséquence du fait que les supports de K_+ et de K_- sont respectivement $[-1,0]$ et $[0,1]$. On a alors, grâce aux propriétés d'indépendance de la suite $(\varepsilon_i)_{i\in\mathbb{N}}$, $\mathbb{E}[(\widehat{m}_+(\tau)-m_+(\tau))(\widehat{m}_-(\tau)-m_-(\tau))|X_1,\ldots,X_n]=B_n^+\,B_n^-$. On retrouve alors le même résultat qu'en indépendant [4]. □

THÉORÈME 2.3. – *On suppose que le processus* **X** *est* $(\theta,\mathcal{L},\psi_1')$ *faiblement dépendant. Si* $\theta_r=r^{-a}$ *avec* $a\geqslant 3$, *et si* $nh^5\to 0$, $nh\to+\infty$ *alors* $\sqrt{nh}(\widehat{\gamma}(\tau)-\gamma(\tau))\xrightarrow{\text{Loi}}\mathcal{N}(0,2\frac{\sigma^2(\tau)}{f(\tau)}V_+)$.

Remarque. – Si le paramètre de fenêtre vérifie $h=n^{-\delta}$, $1/5<\delta<1$, alors le TLC précédent est valable pour $a>6\delta/(1+\delta)$.

Ligne de démonstration. – Le contrôle du terme de biais suit la ligne de démonstration de [4].

$$\sqrt{nh}\big(\mathbb{E}\big(\widehat{m}_+(\tau)|X_1,\ldots,X_n\big)-m_+(\tau)\big)=\sqrt{nh}\frac{\sum_{i=1}^n \omega_i^+(\tau)(m_+(X_i)-m_+(\tau))}{\sum_{i=1}^n \omega_i^+(\tau)}$$
$$=\sqrt{nh}\frac{h^2B_+m_+^{(2)}(\tau)}{2}\big(1+\mathrm{o}_P(1)\big)=\mathcal{O}_P\big(\sqrt{nh^5}\big).$$

On obtient une borne analogue pour $\sqrt{nh}(\mathbb{E}(\widehat{m}_-(\tau)|X_1,\ldots,X_n)-m_-(\tau))$.

On déduit de la décomposition $\widehat{\gamma}(\tau)-\gamma(\tau)=(\widehat{m}_+(\tau)-m_+(\tau))-(\widehat{m}_-(\tau)-m_-(\tau))$, que si $nh^5\to 0$, alors

$$\sqrt{nh}\big(\mathbb{E}\big(\widehat{\gamma}(\tau)|X_1,\ldots,X_n\big)-\gamma(\tau)\big)=\mathrm{o}_P(1). \tag{3}$$

On se propose ensuite de prouver le comportement asymptotique gaussien de $\widehat{\gamma}(\tau)-\mathbb{E}(\widehat{\gamma}(\tau)|X_1,\ldots,X_n)=\sum_{i=1}^n\frac{\Omega_i}{nh^2}$ avec $\Omega_i=\Omega_i^+-\Omega_i^-$, où $\Omega_i^\pm=\frac{(hK_2^\pm-(\tau-X_i)K_1^\pm)K_\pm(\frac{\tau-X_i}{h})\sigma(X_i)\varepsilon_i}{f(\tau)}$.

La technique utilisée est une variation de la méthode de Lindeberg–Rio développée dans [1]. L'idée est d'établir, pour toute fonction $\varphi:\mathbb{R}\to\mathbb{R}$ possédant des dérivées continues et bornées jusqu'à l'ordre trois : $\Delta_n(\varphi)=\mathbb{E}(\varphi(S_n)-\varphi(\sigma_n\eta))\to 0$ avec $S_n=(\sum_{i=1}^n\Omega_i)/\sqrt{nh^3}$, $\sigma_n^2=\mathrm{Var}(S_n)$, et η une variable aléatoire de loi gaussienne centrée réduite. L'hypothèse de causalité, Ψ_1' au lieu de Ψ_1, est essentielle en raison des termes de covariances, non symétriques, considérés dans cette démonstration. □

Références bibliographiques

[1] C. Coulon-Prieur, P. Doukhan, A triangular central limit theorem under a new weak dependence condition, Statist. Probab. Lett. 47 (2000) 61–68.

[2] P. Doukhan, S. Louhichi, A new weak dependence condition and application to moment inequalities, Stochastic Proc. Appl. 84 (1999) 313–343.

[3] J. Fan, I. Gijbels, Variable bandwidth and local linear regression smoothers, Ann. Statist. 20 (1992) 2008–2036.

[4] G. Grégoire, Z. Hamrouni, Change point estimation by local linear smoothing, J. Multivariate Anal. (2001), à paraître.

[5] E. Rio, Théorie asymptotique des processus aléatoires faiblement dépendants, Math. Appl., Vol. 31, Springer-Verlag, 2000.