



Un test non-paramétrique relatif à un modèle de régression avec données dépendantes

Jean Diebolt, Naâmane Laib

► To cite this version:

Jean Diebolt, Naâmane Laib. Un test non-paramétrique relatif à un modèle de régression avec données dépendantes. Annales de l'ISUP, 1993, XXXVII (1-2), pp.3-20. hal-03664817

HAL Id: hal-03664817

<https://hal.science/hal-03664817>

Submitted on 11 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pub. Inst. Stat. Univ. Paris
XXXVII, fasc. 1-2, 1993, 3 à 20

UN TEST NON-PARAMÉTRIQUE RELATIF À UN MODÈLE DE RÉGRESSION AVEC DONNÉES DÉPENDANTES

par

Jean DIEBOLT et Naâmane LAIB

URA CNRS 1321, Université Paris 6, Aile 45-55, 3ème étage,
4, Place Jussieu, 75252 Paris Cedex 05, France.

Résumé. Nous nous intéressons dans ce travail à la définition et aux principales propriétés d'un test non-paramétrique relatif à la fonction de régression d'un modèle de régression non-linéaire, dans le cas où les variables de régression sont dépendantes. Nous démontrons la convergence du processus sur la base duquel est construite la statistique de test. Puis nous étudions la puissance asymptotique locale de ce test.

Abstract. In the present work, we examine the description and main properties of a nonparametric test for the regression function of a nonlinear model for dependent regression variables. We establish the convergence in distribution of the sequence of processes giving rise to the test statistics. Our proof makes use of a weak invariance principle in the α -mixing context. Then, we study the local power of the test.

AMS subject classification: 62M07

1. Introduction

L'objet de ce travail est l'étude d'un test non-paramétrique permettant de tester la fonction de régression T dans le modèle non-paramétrique $Y_n = T(X_n) + \sigma \epsilon_n$. Les deux suites (ϵ_n) et (X_n) sont mutuellement indépendantes et la suite (ϵ_n) est une suite de variables aléatoires réelles i.i.d. centrées et de variance finie. La suite (X_n) est stationnaire et fortement mélangeante (α -mélangeante). On désire tester l'hypothèse nulle $H_0[T = T_0]$ contre l'alternative H_1 , qui peut être composée : $[T \neq T_0]$, ou simple : $[T = T_1]$, sans que l'on connaisse nécessairement la constante σ , ni la loi commune des

variables aléatoires ϵ_n .

Malgré les nombreux travaux relatifs aux modèles de régression non-linéaires, les études de tests paramétriques relatifs à de tels modèles sont peu nombreuses. Parmi les études les plus récentes dans ce domaine, citons celle de Neill [11], qui a étudié un test relatif à un modèle de régression paramétrique non-linéaire multidimensionnel.

Cependant, l'étude d'un test non-paramétrique relatif à un modèle de régression non-linéaire n'a été abordée que dans les travaux de Diebolt [1],[2], où cette procédure de test a été introduite dans le cadre de modèles autorégressifs non-linéaires de séries chronologiques, dans le cas uniformément fortement mélangeant (ϕ -mélangeant), puis Diebolt et Laïb [3], et Laïb [9]. La procédure de test que nous proposons ici est similaire à celle introduite dans [1], [2], et étudiée dans [3],[9]. Néanmoins, le test étudié dans [1],[2] et [3], [9] repose sur l'estimation de la fonction T , donc dépend essentiellement du choix des fenêtres. Par contre, le test que nous proposons ici est indépendant de toute estimation de T .

Notre but, dans cet article, est de démontrer la convergence du processus sur la base duquel est construite la statistique de test, en adaptant la méthode utilisée par Doukhan et Portal [5] pour l'étude du processus empirique dans un cadre α -mélangeant. Puis nous étudions la puissance asymptotique locale du test.

Précisons aussi que le Théorème 1 peut être étendu au cas où les ϵ_n forment une suite dépendante stationnaire, mais que le test proposé devient alors impraticable, car la loi limite de la statistique de test est alors impossible à calculer explicitement, ou à approximer.

De nombreuses questions au sujet de ce test restent posées. Citons l'étude du cas multidimensionnel, l'étude de modèles de régression d'ordre plus grand que 1, et l'étude non-asymptotique des caractéristiques du test.

L'introduction du modèle précédent est motivée par l'étude de la fonction de régression T dans le cas où les X_t sont engendrés par un modèle

autorégressif non-linéaire.

Supposons que $(X_t; t \in \mathbb{N})$ résulte de l'observation d'un processus univarié autorégressif non-linéaire $(X_t; t \in \mathbb{N})$ défini par

$$X_t = \phi(X_{t-1}) + \zeta(X_{t-1})\eta_t, \quad (1.1)$$

où la variable aléatoire X_0 et la suite (η_t) sont indépendantes de la suite (ϵ_t) , ϕ et ζ sont des fonctions réelles, (η_t) est une suite de variables aléatoires i.i.d. centrées et de variance égale à 1. Le processus décrit par (1.1) est une chaîne de Markov homogène. Nous allons montrer que, pour certains modèles classiques, les hypothèses sous lesquelles nous établissons nos résultats sont satisfaites.

Nous commençons par rappeler quelques définitions. Notons $F_n^m = \sigma(X_t, n \leq t \leq n+m)$ la tribu engendrée par $X_n, X_{n+1}, \dots, X_{n+m}$. Le coefficient de mélangeance forte $\alpha(m)$ est défini par

$$\alpha(m) = \sup_{A \in F_0^n, B \in F_{n+m}^\infty} |P(A|B) - P(A)P(B)|.$$

La suite (X_t) est dite fortement mélangeante si $\alpha(m) \rightarrow 0$ quand $m \rightarrow \infty$, et est géométriquement α -mélangeante si $\alpha(m) \leq a_0 \rho^m$, avec $a_0 \geq 0$ et $0 \leq \rho < 1$.

Une chaîne de Markov (X_t) est dite géométriquement ergodique si il existe $0 \leq \rho < 1$ tel que

$$\|P^n(\cdot|x) - \pi(\cdot)\|_{TV} = O(\rho^n) \text{ pour } \pi - \text{presque tout } x,$$

où π est la distribution stationnaire de (X_t) et $\|\cdot\|_{TV}$ la norme de la variation totale, et où P^n est la n^e probabilité de transition de (1.1) (voir, par exemple, Nummelin et Tuominen (1982, [12])). Faisons les hypothèses suivantes :

(H1) Pour tout compact I de $\mathbb{R}$, il existe un réel $b > 0$ tel que $\zeta^2(x) \geq b$ pour tout $x \in I$.

(H2) Pour tout compact I de $\mathbb{R}$, $\inf_{x \in I} \mu_{\eta_0}(x) > 0$, où μ_{η_0} est la densité de la variable aléatoire η_0 .

1) Modèles à seuils.

Soit $\{I_j\}_{j \in J}$ une partition de $\mathbf{R}$. Si, pour $x \in I_j$, $\phi(x) = a(j)x + c(j)$, et que $\zeta(x)$ est constante et égale à $\sigma(j) > 0$, (1.1) correspond au modèle à seuils SETAR introduit et étudié par Tong et Lim (1980). Sous (H1) et (H2) et si de plus $|a(j)| < 1$ pour tout j , alors (X_t) est géométriquement ergodique et α -mélangeante avec une vitesse de convergence géométrique (voir Diebolt et Guégan [4]).

2) Modèle EXPAR.

Si $\zeta(x)$ est constante et égale à $\sigma_0 > 0$ et $\phi(x) = (a + b \exp(-cx^2))x$, où a, b, c sont des constantes et $c > 0$, on obtient le modèle EXPAR introduit par Ozaki (1979). Sous (H1) et (H2) et si de plus $|a| < 1$, (X_t) est géométriquement ergodique et α -mélangeante avec une vitesse de convergence géométrique (voir Diebolt et Guégan [4]).

3) Modèle bilinéaire.

Si $\phi(x) = ax$, $\zeta(x) = bx + \sigma_0$, où σ_0 est une constante strictement positive, on obtient le modèle markovien bilinéaire. Sous (H1) et (H2) et si de plus $|a| < 1$, $a^2 + b^2 < 1$, (X_t) est géométriquement ergodique et α -mélangeante avec une vitesse de convergence géométrique (voir Diebolt et Guégan [4]).

4) Modèle ARCH.

Si $\phi(x) = rx$, $\zeta(x) = \sqrt{a_0 + a_1 x^2}$, où a_0 et a_1 sont des constantes positives, on obtient le modèle AR univarié avec bruit ARCH introduit et étudié par Engle (1982) dans le cas où $r = 0$ et par Weiss (1984) dans le cas $r \neq 0$. Sous (H1) et (H2) et si $|r| < 1$, $0 \leq a_1 < 1$, $a_0 > 0$, $|r| + a_1^{\frac{1}{2}} E(|\eta_0|) < 1$, (X_t) est géométriquement ergodique et α -mélangeante avec une vitesse de convergence géométrique (voir Diebolt et Guégan [4]).

5) Modèle β -ARCH.

Si $\phi(x) = rx$, $\zeta(x) = \sqrt{a_0 + a_1 x^{2\beta}}$, on obtient le modèle AR univarié avec bruit β -ARCH. Sous (H1) et (H2) et si $a_0, a_1 > 0$, $0 \leq \beta < 1$, alors (X_t)

est géométriquement ergodique et α -mélangeante avec une vitesse de convergence géométrique si $|r| < 1$ (voir Diebolt et Guégan [4]).

2. Description du test

Nous considérons des suites de variables aléatoires $(X_n; n \geq 1)$, $(Y_n; n \geq 1)$ et $(\epsilon_n; n \geq 1)$ reliées par $Y_n = T(X_n) + \sigma\epsilon_n$. La suite $(X_n; n \geq 1)$ est supposée stationnaire et α -mélangeante, la série $\sum_{n \geq 1} \alpha(n)^\gamma$ étant convergente pour un $0 < \gamma < 1$, la suite $(\epsilon_n; n \geq 0)$ est supposée i.i.d. de moyenne 0 et de variance 1, et on suppose que $E(|\epsilon_0|^4) < \infty$. Les deux suites (X_n) et (ϵ_n) sont supposées mutuellement indépendantes. La fonction T est continue sur $\mathbf{R}$ et σ est strictement positive. Enfin, on suppose que la fonction de répartition F de la loi de X_1 est continue et strictement croissante.

On définit le processus suivant sous l'hypothèse (H_0) :

$$A_n(t) = n^{-\frac{1}{2}} \sum_{i=1}^n \{Y_i - T_0(X_i)\} 1_{(-\infty, t]}(X_i), \quad t \in \mathbf{R}. \quad (2.1)$$

La statistique de test choisie ici est $\sup_{t \in \mathbf{R}} |A_n(t)|$, que l'on notera M_n . On prouve au paragraphe 3 que, sous des conditions très générales, M_n converge en distribution vers $\sigma \sup_{t \in [0,1]} |W(t)|$, où W est un processus de Wiener sur $\mathbf{R}^+$. Pour tout niveau de signification approché α , $0 < \alpha < 1$, la région critique du test est

$$R_n = \{x_1, \dots, x_n : M_n > q_\alpha\}, \quad (2.2)$$

où q_α est le quantile d'ordre $1 - \alpha$ de la loi limite de M_n : si l'on pose $\Psi(x) = P(\max_{t \in [0,1]} |W(t)| \leq x)$, $x \geq 0$, et si l'on choisit x_α tel que $1 - \Psi(x_\alpha) = \alpha$, alors

$$q_\alpha = \sigma x_\alpha. \quad (2.3)$$

Comme la quantité σ est inconnue, on doit la remplacer par un estimateur σ_n p.s. convergent sous H_0 , défini comme suit :

$$\sigma_n^2 = n^{-1} \sum_{i=1}^n \{Y_i - T_0(X_i)\}^2.$$

La région critique devient alors

$$R'_n = \{x_1, \dots, x_n : M_n > q_{n,\alpha}\}, \quad (2.4)$$

où

$$q_{n,\alpha} = \sigma_n x_\alpha. \quad (2.5)$$

3. Résultats de convergence sous H_0

Nous notons $\|\cdot\|$ la norme de la convergence uniforme sur $\mathbf{R}$. Pour établir sous H_0 la convergence de A_n vers le processus gaussien $A \stackrel{d}{=} \sigma(W \circ F)$, on établit un principe d'invariance faible analogue à celui de Doukhan et Portal [5] pour le processus empirique.

Pour étudier la convergence de la suite (A_n) , il est commode d'introduire la suite des processus auxiliaires B_n , définis sur $\mathbf{R}$ par

$$B_n(t) = \sigma n^{-\frac{1}{2}} \sum_{k=1}^n g_{n,t}(\xi_k), \quad (3.1)$$

où $\xi_k = (X_k, \epsilon_k)$ et $g_{n,t}(\xi_k) = 1_{(-\infty, t]}(X_k) \epsilon_k$. Comme F est supposée continue et strictement croissante, son inverse F^{-1} est continue et strictement croissante. Posons $V_i = F(X_i)$; alors $(V_i; i \geq 1)$ est une suite de v.a. uniformes sur $[0, 1]$, stationnaire et α -mélangeante avec le même coefficient de mélangeance que $(X_i; i \geq 1)$.

La définition (3.1) devient

$$B_n^*(s) = \sigma n^{-\frac{1}{2}} \sum_{k=1}^n g_{n,s}^*(\xi_k^*) \quad s \in [0, 1], \quad (3.2)$$

où $\xi_k^* = (V_k, \epsilon_k)$ et $g_{n,s}^*(\xi_k^*) = 1_{(0,s]}(V_k) \epsilon_k$, et où $B_n^* = B_n \circ F^{-1}$. Sous l'hypothèse nulle H_0 , les deux suites (A_n) et (B_n) ont même limite, ainsi que (A_n^*) et (B_n^*) , avec $A_n^* = A_n \circ F^{-1}$.

THEOREME 1. Soit (ξ_k) la suite α -mélangeante de variables aléatoires bivariées (X_k, ϵ_k) , supposée définie sur un espace probabilisé $(\Omega^*, \Sigma^*, P^*)$ suffisamment riche pour que l'on puisse y définir une variable aléatoire uniforme

sur $[0, 1]$, indépendante des ξ_k . Alors, il existe une suite $(\tilde{Y}_n; n \geq 1)$ de processus gaussiens centrés, de fonction de covariance $\sigma^2 F(s \wedge t)$, $s, t \in \mathbf{R}$, telle que

$$\|B_n - \tilde{Y}_n\| \xrightarrow{P} 0. \quad (3.3)$$

Nous en déduisons clairement les corollaires suivants.

COROLLAIRE 1. $\|A_n - \tilde{Y}_n\| \xrightarrow{P} 0.$

COROLLAIRE 2. $|\sup_{t \in \mathbf{R}} |A_n(t)| - \sup_{t \in \mathbf{R}} |\tilde{Y}_n(t)|| \xrightarrow{P} 0.$

COROLLAIRE 3. La suite de processus $(A_n; n \geq 1)$ converge en distribution vers le processus gaussien $A = \sigma(W \circ F)$.

COROLLAIRE 4. $M_n \xrightarrow{d} \sigma \sup_{t \in \mathbf{R}} |W(F(t))|.$

Principe de la preuve du Théorème 1 et préliminaires.

On se ramène d'abord à l'étude des processus B_n^* . Les variables aléatoires considérées ici prennent leurs valeurs dans l'espace de Skorohod $E = D[0, 1]$ des fonctions x définies sur $[0, 1]$, continues à droite et admettant une limite à gauche en tout point de $[0, 1]$. On définit, à une équivalence près, la distance de Skorohod entre deux éléments x et y de E par

$$d(x, y) = \inf \{ \epsilon > 0 : (\exists \lambda \in \Lambda_{0,1}) \|\lambda\| \leq \epsilon, \|x - y \circ \lambda\| \leq \epsilon \}, \quad (3.4)$$

où $\Lambda_{0,1} = \{ \lambda : [0, 1] \rightarrow [0, 1] : \lambda(0) = 0, \lambda(1) = 1 \text{ et } \lambda \text{ est continue et strictement croissante} \}$. On note $(E, \mathfrak{F}, d)$ l'espace métrique mesurable muni de la distance d , où $\mathfrak{F}$ désigne l'ensemble des boréliens de E ; on note aussi $\omega_x(\delta), \delta > 0$, le module de continuité de la fonction x . La distance de Prohorov relative à la distance d entre deux lois de probabilités P et Q sur l'espace E est définie par

$$\rho_d(P, Q) = \inf \{ \epsilon > 0 : (\forall A \in F_E) P(A) \leq Q(A^c) + \epsilon, Q(A) \leq P(A^c) + \epsilon \},$$

où l'on a noté $A^c = \{ x \in E : d(x, A) \leq \epsilon \}$ et F_E l'ensemble des fermés de E .

Les théorèmes de Strassen [14] et de Dudley et Philipp [6] permettent de ramener notre problème à une estimation de la distance de Prohorov ρ_d entre lois de processus.

Notre évaluation de la distance $\rho_d(P_n, Q_n)$, où P_n est la loi du processus B_n^* et Q_n celle d'un processus gaussien centré de covariance $\sigma^2(s \wedge t)$, repose sur le lemme suivant.

LEMME A (Doukhan et Portal (1987, [5])). Soient X et Y deux variables aléatoires à valeurs dans E et $K = \{t_1, \dots, t_k\}$ un δ -réseau. Alors,

$$\rho_d(P_X, P_Y) \leq \epsilon_X(\delta) + \epsilon_Y(\delta) + \rho_{\|\cdot\|}(P_X \circ \Pi_K^{-1}, P_Y \circ \Pi_K^{-1}), \quad (3.5)$$

où $\rho_{\|\cdot\|}$ désigne la distance de Prohorov relative à la norme sup sur $\mathbf{R}^k$, Π_K est la projection canonique de E sur $\mathbf{R}^k$ associée à K ,

$$\epsilon_X(\delta) = \inf\{\epsilon > 0 : P\{\omega_X(\delta) > \epsilon\} < \epsilon\},$$

où ω_X désigne le module de continuité de X , P_X la loi de X et $P_X \circ \Pi_K^{-1}$ celle de $(X(t_1), \dots, X(t_k))$.

Dans tout ce qui suit, on prend $X = B_n^*$. Par le Lemme A, on détermine une suite $(\alpha_n; n \geq 1)$ tendant vers 0, telle que $\rho_d(P_n, Q_n) \leq \alpha_n$, et le théorème de Strassen [14] implique alors l'existence d'une loi de probabilité F_n sur $(E^2, \mathfrak{F}^2)$, de marginales P_n et Q_n , telle que

$$F_n\{(x, y) \in E^2 : d(x, y) > \alpha_n\} < \alpha_n.$$

Enfin, le théorème de Dudley et Philipp [6] nous permet de construire sur $(\Omega^*, \Sigma^*, P^*)$ des processus gaussiens centrés Y_n^* de covariance $\sigma^2(s \wedge t)$, tels que

$$P\{d(B_n^*, Y_n^*) > \alpha_n\} < \alpha_n. \quad (3.6)$$

Preuve du Théorème 1. Les points essentiels qui restent à étudier sont les oscillations des processus B_n^* et des processus gaussiens de loi Q_n , puis leurs répartitions fini-dimensionnelles. Commençons par l'étude des oscillations des processus B_n^* .

LEMME 1. Supposons que le coefficient de mélangeance de la suite $(V_n; n \geq 1)$ vérifie $\sum_{n=1}^{\infty} \alpha(n)^{\gamma} < \infty$, pour un réel γ tel que $0 < \gamma < 1$. Soit

$u_n(\delta) = d_{\alpha, \gamma}^{\frac{1}{2}} (n^{-1} \delta^{-\gamma} + \delta)^{\frac{1}{2}}$, où $d_{\alpha, \gamma} = C_1 \max(1, k(\alpha, \gamma)) \times \max(1, 2^{\gamma})$ et où $k(\alpha, \gamma) = 48 \sum_{n=1}^{\infty} \alpha(n)^{\gamma} < \infty$ et C_1 est une constante qui provient de l'inégalité de Von Bahr-Esseen. Alors, le processus B_n^* vérifie

$$P\{\omega_{B_n^*}(\delta) \geq u_n(\delta)\} < u_n(\delta). \quad (3.7)$$

Preuve. Nous considérons la statistique ordonnée $V_{1:n} < V_{2:n} < \dots < V_{n:n}$ de l'échantillon $\{V_1, \dots, V_n\}$. Il existe une suite de variables aléatoires (ϵ_n^*) , définie sur le même espace de probabilité que (ϵ_n) , et identique en loi à (ϵ_n) , telle que

$$B_n^* \stackrel{d}{=} \sigma(S_n \circ G_n), \quad (3.8)$$

où S_n est le processus des sommes partielles de (ϵ_n^*) , défini par

$$S_n(u) = 0 \text{ pour } 0 \leq u \leq \frac{1}{n},$$

$$S_n(u) = \frac{S_j}{\sqrt{n}} \text{ pour } \frac{j}{n} \leq u < \frac{j+1}{n} \text{ et } 1 \leq j \leq n,$$

avec $S_j = \sum_{i=1}^j \epsilon_i^*$, et $G_n(t)$ est la fonction de répartition empirique de l'échantillon $\{V_1, \dots, V_n\}$.

Soit $0 \leq \delta \leq 1$ et M l'entier naturel donné par $M = [\frac{1}{\delta}] + 1$ (où $[t]$ désigne la partie entière de t). On divise l'intervalle $[0, 1]$ en M sous-intervalles de longueur égale à $\frac{1}{M}$. Par conséquent,

$$\begin{aligned} \omega_{B_n^*}(\delta) &= \sup_{0 \leq t-s \leq \delta} |B_n^*(t) - B_n^*(s)| \\ &\leq 3 \max_{1 \leq j \leq M-1} \{ \sup_{0 \leq t \leq \frac{1}{M}} |B_n^*(\frac{j}{M} + t) - B_n^*(\frac{j}{M})| \}. \end{aligned}$$

Comme la suite B_n^* est à accroissements stationnaires, on obtient, pour tout $\lambda > 0$,

$$\begin{aligned} P\{\omega_{B_n^*}(\delta) \geq \lambda\} &\leq \sum_{j=1}^M P\{\|B_n^*\|_0^{\frac{1}{M}} \geq \frac{\lambda}{3}\} \\ &\leq M P\{\|B_n^*\|_0^{\frac{1}{M}} \geq \frac{\lambda}{3}\}, \end{aligned} \quad (3.9)$$

où $\|B_n^*\|_0^{\frac{1}{M}} = \sup_{0 \leq t \leq \frac{1}{M}} |B_n^*(t)|$. Comme $G_n(t)$ est croissante, on obtient d'après (3.8)

$$P\{\omega_{B_n^*}(\delta) \geq \lambda\} \leq MP\{\sigma \max_{1 \leq j \leq nG_n(1/M)} \left| \frac{S_j}{\sqrt{n}} \right| \geq \frac{\lambda}{3}\}.$$

Notons $\mathbf{V} = \{V_1, \dots, V_n\}$ et $\mathbf{v} = \{v_1, \dots, v_n\}$, et posons

$$K_{n,M} = \sum_{i=1}^n 1_{(V_i \leq \frac{1}{M})} = nG_n\left(\frac{1}{M}\right).$$

Par l'inégalité de Markov, on a

$$P\{\omega_{B_n^*}(\delta) \geq \lambda\} \leq 81\sigma^4 ME(\max_{1 \leq j \leq K_{n,M}} \left| \frac{S_j}{\sqrt{n}} \right|^4) / \lambda^4. \quad (3.10)$$

On commence par évaluer le membre de droite de (3.10) conditionnellement à $(\mathbf{V} = \mathbf{v})$. En tenant compte d'une inégalité de Von Bahr-Esseen (voir Shorack et Wellner [13], page 858), on obtient

$$E(\max_{1 \leq i \leq K_{n,M}} \left| \frac{S_i}{\sqrt{n}} \right|^4 | \mathbf{V} = \mathbf{v}) \leq \frac{8}{n^2} E(S_{K_{n,M}}^4 | \mathbf{V} = \mathbf{v}).$$

Or, si l'on choisit dans l'inégalité 2 de Von Bahr-Esseen (voir Shorack et Wellner [13], page 858) $r = 4$, on obtient

$$E(S_{K_{n,M}}^4 | \mathbf{V} = \mathbf{v}) \leq 2C_1 E(K_{n,M}^2 | \mathbf{V} = \mathbf{v}), \quad (3.11)$$

où C_1 est une constante qui provient de l'inégalité de Von Bahr-Esseen. D'où

$$P\{\omega_{B_n^*}(\delta) \geq \lambda\} \leq \frac{C_2 \sigma^4 M}{n^2} E(K_{n,M}^2) / \lambda^4. \quad (3.12)$$

Pour l'évaluation du terme $E(K_{n,M}^2)$, posons $\tilde{K}_{n,M} = \sum_{i=1}^n Z_i$, où $Z_i = 1_{(V_i \leq \frac{1}{M})} - \frac{1}{M}$. Donc, on a $K_{n,M} = \tilde{K}_{n,M} + \frac{n}{M}$, et, comme $\tilde{K}_{n,M}$ est une variable aléatoire centrée, on a

$$E(K_{n,M}^2) = E(\tilde{K}_{n,M}^2) + \left(\frac{n}{M}\right)^2. \quad (3.13)$$

Pour évaluer le membre de droite de (3.13), on applique un résultat de Doukhan et Portal [5], Théorème 3, aux variables aléatoires $Z_i = 1_{(V_i \leq \frac{1}{M})} - \frac{1}{M}$.

Or, pour tout $i \in N$, $|Z_i| \leq 1$ et $E K_{n,M} = 0$. Par suite, pour $0 < \gamma < 1$, on a

$$E(|Z_i|^{\frac{2}{1-\gamma}}) \leq E(|Z_i|^2) \leq \frac{1}{M}$$

et

$$(E(|Z_i|^{\frac{2}{1-\gamma}}))^{1-\gamma} \leq \frac{1}{M^{1-\gamma}}.$$

Donc

$$E(K_{n,M}^2) \leq k(\alpha, \gamma) n M^{-1+\gamma} + \left(\frac{n}{M}\right)^2, \quad (3.14)$$

où $k(\alpha, \gamma) = 48 \sum_{i \geq 1} \alpha_i^\gamma$, et

$$P\{\omega_{B_n^\bullet}(\delta) \geq \lambda\} \leq C_3(\alpha, \gamma)[n^{-1}M^\gamma + M^{-1}]/\lambda^4, \quad (3.15)$$

où $C_3(\alpha, \gamma) = C_2 k'(\alpha, \gamma)$ avec $k'(\alpha, \gamma) = \max(1, k(\alpha, \gamma))$. Comme M est choisi de telle sorte que $\frac{1}{\delta} \leq M \leq \frac{2}{\delta}$, ceci implique

$$P\{\omega_{B_n^\bullet}(\delta) \geq \lambda\} \leq C_4(\alpha, \gamma)[n^{-1}\delta^{-\gamma} + \delta]/\lambda^4, \quad (3.16)$$

où $C_4(\alpha, \gamma) = C_3(\alpha, \gamma) \max(1, 2^\gamma)$. Si l'on pose

$$u_n(\delta) = [C_4(\alpha, \gamma)(n^{-1}\delta^{-\gamma} + \delta)]^{\frac{1}{5}},$$

on obtient

$$P\{\omega_{B_n^\bullet}(\delta) \geq u_n(\delta)\} \leq u_n(\delta). \quad \square \quad (3.17)$$

L'étude des oscillations des processus gaussiens de loi Q_n se ramène à celle des oscillations du processus de Wiener.

Pour les répartitions fini-dimensionnelles, on remarque tout d'abord que, si $|x|$ désigne la norme euclidienne de $\mathbf{R}^k$, alors $\rho_{\|\cdot\|} \leq \rho_{|\cdot|}$ (voir [7]). Donc, il suffit de contrôler $\rho_{|\cdot|}(P_n \circ \Pi_K^{-1}, Q_n \circ \Pi_K^{-1})$. Pour évaluer ce dernier terme, nous utilisons un résultat de Dehling [7].

Posons, dans le Théorème 3 de Dehling [7], $B = \mathbf{R}^k$, $K = K_n = \{t_1, \dots, t_{k(n)}\}$ étant une suite de δ -réseaux de $[0, 1]$ telle que $\delta = \delta_n \rightarrow 0$ quand $n \rightarrow \infty$, $\xi_i^* =$

$\sigma 1_{[0,s]}(V_l)\epsilon_i$; alors, $\frac{S_n}{\sqrt{n}} = \frac{1}{\sqrt{n}}(B_n^*(t_1), \dots, B_n^*(t_k))^t$, $\nu = P_n o \Pi_K^{-1}$, $\mu = Q_n o \Pi_K^{-1}$. On a donc

$$\rho_{1,1}(P_n o \Pi_K^{-1}, Q_n o \Pi_K^{-1}) = O(n^{-\tau} k^3(n)), \quad (3.18)$$

où $\tau = \frac{2\theta\beta}{200(2+\beta)}$, avec $0 < \theta < \frac{2}{3}$ et $0 < \beta \leq 1$, si le coefficient de mélangeance $\alpha(n)$ vérifie $\alpha(n) \leq n^{-(1+\beta)(1+\frac{\theta}{2})}$. Si l'on choisit des δ_n -réseaux équidistribués, $\beta = 1$ et $\theta = \frac{1}{2}$, alors $k(n)\delta_n = 1$ et

$$\rho_d(P_n, Q_n) \leq \epsilon_{B_n^*}(\delta_n) + \epsilon_n(\delta_n) + O(n^{-\frac{1}{600}} \delta_n^{-3}). \quad (3.19)$$

Or, le Lemme 1 implique que $\epsilon_{B_n^*}(\delta_n) = O((n^{-1} \delta_n^{-\gamma} + \delta_n)^{\frac{1}{2}})$, un résultat de Bickel et Freedman pour le module de continuité du processus de Wiener (voir Schorack et Wellner [13], pages 538-540) implique que

$\epsilon_n(\delta_n) = O(\delta_n \log(\frac{1}{\delta_n})^{\frac{1}{4}})$. Si l'on choisit δ_n de manière appropriée, on obtient

$$\rho_d(P_n, Q_n) = O(\alpha_n), \quad (3.20)$$

où $\alpha_n \rightarrow 0$ quand $n \rightarrow \infty$. Les théorèmes de Strassen et de Dudley et Philipp nous permettent alors de construire des versions des processus gaussiens centrés Y_n^* de covariance $\sigma^2(s \wedge t)$, tels que

$$P\{d(B_n^*, Y_n^*) > \alpha_n\} < \alpha_n. \quad (3.21)$$

Or, pour tous x, y dans E , $\|x - y\| \leq d(x, y) + \omega_x[d(x, y)]$ (voir [5]), ceci implique que $P\{\|B_n^* - Y_n^*\| \geq 2\alpha_n\} \leq 2\alpha_n$ dès que $P(\omega_n(\alpha_n) > \alpha_n) < \alpha_n$, car

$$P\{\|B_n^* - Y_n^*\| \geq 2\alpha_n\} \leq \alpha_n + P\{\|B_n^* - Y_n^*\| > 2\alpha_n \text{ et } d(B_n^*, Y_n^*) \leq \alpha_n\}$$

et que, sur l'évènement $\{\|B_n^* - Y_n^*\| > 2\alpha_n \text{ et } d(B_n^*, Y_n^*) \leq \alpha_n\}$, on a

$$\|B_n^* - Y_n^*\| \leq \alpha_n + \omega_n(\alpha_n).$$

Par conséquent,

$$P\{\|B_n^* - Y_n^*\| > 2\alpha_n \text{ et } d(B_n^*, Y_n^*) \leq \alpha_n\} \leq P\{\omega_n(\alpha_n) > \alpha_n\},$$

d'où

$$P\{\|B_n^* - Y_n^*\| \geq 2\alpha_n\} \leq 2\alpha_n. \quad \square$$

4. Etude locale de la puissance du test

Soit T la fonction à tester sur $\mathbf{R}$. Posons $\Delta T = (T - T_0)$, et soit F la fonction de répartition commune des variables aléatoires $X_1, \dots, X_n$. Nous établissons dans ce paragraphe un développement asymptotique du processus A_n . Le résultat obtenu permet de préciser le comportement asymptotique de l'erreur de deuxième espèce lorsque ΔT est suffisamment petit ce qui correspond à des hypothèses H_0 et H_1 dites contigues.

THÉORÈME 2. *Soit (ξ_k) la suite de variables aléatoires bivariées (X_k, ϵ_k) , où $(X_k; k \geq 1)$ est strictement stationnaire et α -mélangeante, et où le coefficient de mélangeance $\alpha(n)$ vérifie, pour tout entier l ,*

$$\lim_{n \rightarrow \infty} n^l \alpha(n) = 0. \quad (4.1)$$

La suite (ξ_k) est supposée définie sur l'espace probabilisé $(\Omega^, \Sigma^*, P^*)$. Si ΔT est mesurable et bornée et si $\int_{-\infty}^{\infty} d|\Delta T| < \infty$, alors, pour n assez grand, le processus $A_n(t)$ peut s'écrire*

$$A_n(t) = n^{\frac{1}{2}} \int_{-\infty}^t \Delta T dF + \sigma W_n(F(t)) + R_{n,3}(t), \quad (4.2)$$

où W_n est la suite de processus de Wiener définis dans le Théorème 1, et

$$\|R_{n,3}\| = O_P((\log n)^{\frac{1}{2}}).$$

Preuve du Théorème 2. Le processus empirique de l'échantillon $\{V_i; i = 1, \dots, n\}$ sera noté

$$U_n = n^{\frac{1}{2}} \{G_n - I\}, \quad (4.3)$$

où I désigne la fonction identité ($I(x) = x$). Nous utiliserons aussi la notation $X_n = O_P(h_n)$ qui signifie : pour tout $\epsilon > 0$, il existe un nombre réel M_ϵ tel que $P\{|X_n| \geq M_\epsilon h_n\} \leq \epsilon$ pour tout n .

Nous utiliserons, pour prouver le Théorème 2, le Théorème 1 ainsi qu'un

résultat de [5].

Le processus A_n défini par (2.2) peut s'écrire comme suit :

$$A_n(t) = B_n(t) + P_n(t), \quad (4.4)$$

où B_n est le processus défini par (3.1),

$$P_n(t) = n^{-\frac{1}{2}} \sum_{k=1}^n \Delta T(X_k) 1_{(-\infty, t]}(X_k), \quad (4.5)$$

Examinons le comportement asymptotique de chacun de ces termes sous H_1 (donc, avec $\Delta T \neq 0$). On sait déjà que $B_n(t) = \sigma W_n(F(t)) + R_{n,1}(t)$, avec $P\{|R_{n,1}| \geq 2\alpha_n\} \leq 2\alpha_n$ (Théorème 1).

Pour étudier le comportement de $P_n(t)$, on remarque que

$$\begin{aligned} n^{-\frac{1}{2}} P_n(t) &= n^{-1} \sum_{k=1}^n \Delta T(X_k) 1_{(-\infty, t]}(X_k) \\ &= n^{-1} \sum_{k=1}^n (\Delta T \circ F^{-1})(V_k) 1_{[0, F(t)]}(V_k) \\ &= \int_0^s (\Delta T \circ F^{-1}) dG_n, \quad s = F(t). \end{aligned}$$

LEMME 2. Si ΔT est bornée et si $\int_{-\infty}^{\infty} d|\Delta T| < \infty$, alors le processus P_n défini par (4.5) vérifie, quand $n \rightarrow \infty$,

$$P_n(t) = n^{\frac{1}{2}} \int_{-\infty}^t \Delta T dF + R_{n,2}(t), \quad (4.6)$$

où $\|R_{n,2}\| = O_P((\log n)^{\frac{1}{2}})$.

Preuve du Lemme 2. On a

$$P_n(t) = n^{\frac{1}{2}} \int_0^s (\Delta T \circ F^{-1}) dI + \int_0^s (\Delta T \circ F^{-1}) dU_n.$$

D'où

$$P_n(t) = n^{\frac{1}{2}} \int_0^s (\Delta T \circ F^{-1}) dI + R_{n,2}(t),$$

avec $s = F(t)$ et où l'on a posé

$$R_{n,2}(s) = (\Delta T o F^{-1})(s) U_n(s) - \int_0^s U_n d(\Delta T o F^{-1}).$$

Si l'on pose $C_5 = \|\Delta T o F^{-1}\| \int_{-\infty}^{\infty} d|\Delta T|$, on obtient

$$\|R_{n,2}\| \leq C_5 \|U_n\|.$$

Or, si le coefficient de mélangeance vérifie (4.1), il existe une suite $(\tilde{U}_n; n \geq 1)$ de processus gaussiens centrés de même loi, telle que si l'on choisit $\lambda_n \geq 2b_1 C_5 n^{-\frac{1}{27}}$, on ait

$$P\{\|U_n - \tilde{U}_n\| \geq \frac{\lambda_n}{2C_5}\} = O(n^{-\frac{1}{27}}),$$

où b_1 est une constante qui provient du Théorème 1 de Doukhan et Portal [5], et, pour n fixé, la suite $(\tilde{U}_n)$ est une suite de variables aléatoires gaussiennes de même loi à valeurs dans l'espace de Banach des fonctions continues de $\mathbf{R}$ dans $\mathbf{R}$, muni de la norme de la convergence uniforme. Par le Théorème d'intégrabilité de Fernique, il existe une constante $b_2 > 0$ telle que l'on ait, pour tout entier $n \geq 1$:

$$P\{\|\tilde{U}_n\| > \lambda\} = O(\exp(-b_2 \lambda^2)) \text{ quand } \lambda \rightarrow \infty$$

(voir, par exemple, Sato [10]). Par conséquent,

$$\begin{aligned} P\{\|R_{n,2}\| \geq (\log n)^{\frac{1}{2}}\} &\leq P\{\|U_n - \tilde{U}_n\| \geq \frac{\lambda_n}{2C_5}\} + P\{\|\tilde{U}_n\| \geq \frac{(\log n)^{\frac{1}{2}}}{2C_5}\} \\ &\leq O(n^{-\frac{1}{27}}) + O(n^{-\frac{b_2}{4C_5^2}}) = O(\beta_n), \end{aligned}$$

où $\beta_n \rightarrow 0$ quand $n \rightarrow \infty$. Si l'on pose $R_{n,3}(t) = R_{n,1}(t) + R_{n,2}(t)$, on obtient

$$\begin{aligned} P\{\|R_{n,3}\| \geq (\log n)^{\frac{1}{2}}\} &\leq P\{\|R_{n,1}\| \geq 2\alpha_n\} + P\{\|R_{n,2}\| \geq (\log n)^{\frac{1}{2}}\} \\ &\leq O(\alpha_n) + O(\beta_n). \end{aligned}$$

D'où le résultat. $\square$

Nous passons maintenant à l'étude directe de la puissance locale du test.

Pour cela, posons $\Delta T(x) = \frac{g(x)}{\sqrt{n}}$ où g est une fonction à valeurs réelles.

THÉORÈME 3. Soit $\nu = \exp(-\frac{1}{2\sigma^2} \int_{-\infty}^{\infty} g^2 dF)$ et supposons que g est une fonction bornée. Alors, pour tout niveau de signification approché α , $0 < \alpha < 1$, on a :

$$\nu(1 - \alpha) \leq \lim_{n \rightarrow \infty} P\{\sup_{t \in \mathbb{R}} |A_n(t)| \leq q_{n,\alpha}\} \leq 1 - \alpha.$$

Preuve du théorème 3. Si $\Delta T(x) = \frac{g(x)}{\sqrt{n}}$, alors (4.2) devient

$$A_n(t) = \int_{-\infty}^t g dF + \sigma W_n(F(t)) + R_{n,4}(t), \quad (4.7)$$

où $\|R_{n,4}\| = O_P(\alpha_n) + O_P(n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}) \rightarrow 0$ quand $n \rightarrow \infty$ (d'après le Théorème 1 et le Lemme 2). Notons $q_{n,\alpha}$ le quantile d'ordre $(1 - \alpha)$ de la loi exacte de M_n . On sait, d'après le Théorème 1, que $q_{n,\alpha} \rightarrow q_\alpha$ quand $n \rightarrow \infty$ (q_α est le quantile d'ordre $1 - \alpha$ de la loi limite de M_n). Le risque d'une erreur de deuxième espèce converge quand $n \rightarrow \infty$, en vertu de (4.7), vers

$$P\{\sup_{t \in \mathbb{R}} |\int_{-\infty}^t g dF + \sigma W(F(t))| \leq q_\alpha\}. \quad (4.8)$$

Si l'on pose $s = F(t)$ la probabilité (4.8) devient

$$P\{\sup_{s \in [0,1]} |W(s) - a(s)| \leq q_\alpha \sigma^{-1}\}, \quad (4.9)$$

où $a(s) = -\int_0^s g(F^{-1}(\sigma x)) dx$. Or, pour toute fonction f absolument continue par rapport à la mesure de Lebesgue et telle que $\int_0^1 (f'(x))^2 dx < \infty$, on a, pour tout $\lambda > 0$,

$$\Psi(\lambda) \exp(-\frac{1}{2} \int_0^1 (f'(x))^2 dx) \leq P\{\sup_{0 \leq t \leq 1} |W(t) - f(t)| \leq \lambda\} \leq \Psi(\lambda),$$

(voir Gill [8]). Si l'on pose $\nu = \exp(-\frac{1}{2} \int_0^1 (a'(s))^2 ds) = \exp(-\frac{1}{2\sigma^2} \int_{-\infty}^{\infty} g^2 dF)$, $\lambda = \sigma^{-1} q_\alpha$, on a, d'après (2.3), $(1 - \Psi)(\sigma^{-1} q_\alpha) = \alpha$. Par conséquent, la probabilité (4.9) devient

$$\nu(1 - \alpha) \leq P\{\sup_{0 \leq s \leq 1} |W(s) - a(s)| \leq q_\alpha \sigma^{-1}\} \leq 1 - \alpha. \quad \square$$

References

- [1] Diebolt, J. (1985) Deux tests non-paramétriques pour un modèle autorégressif non linéaire, *C. R. Acad. Sci. Paris I*, **301**, 649-652.
- [2] Diebolt, J. (1990) Testing the functions defining a nonlinear autoregressive time series, *Stoch. Proc. and Appl.*, **36**, 85-106.
- [3] Diebolt, J. et Laïb, N. (1991) Un principe d'invariance faible pour l'étude d'un test non-paramétrique relatif à un modèle de régression non-linéaire, *C. R. Acad. Sci. Paris I*, **312**, 887-891.
- [4] Diebolt, J. et Guégan, D. (1990) Probabilistic properties of the general nonlinear Markovian Process of order one and applications to time series modelling, *Rapport Technique L.S.T.A N° 126, Université Paris 6*.
- [5] Doukhan, P. et Portal, F. (1987) Principe d'invariance faible pour la fonction de répartition empirique dans un cadre multidimensionnel et fortement mélangeant, *Probab. Math. Statist.*, **8**, 117-132.
- [6] Dudley, R. et Philipp, W. (1983) Invariance principles for sums of Banach space valued random elements and empirical processes, *Z.W.G.*, **62**, 509-552.
- [7] Dehling, H. (1983) Limit theorems for sums of weakly dependent Banach space valued random variables, *Z.W.G.*, **63**, 393-432.
- [8] Grill, K. (1991) A lim inf result in Strassen's law of the iterated logarithm, *Probab. Th. Rel. Fields*, **89**, 149-157.
- [9] Laïb, N. (1991) Étude de la puissance d'un test non-paramétrique relatif à un modèle de régression non-linéaire, *C. R. Acad. Sci. Paris I*, **313**, 527-532.
- [10] Marcus, M.B. et Shepp, L.A. (1972) Sample behavior of Gaussian processes, Sixth Berkeley symposium on Mathematical Statistics and Probability, **II**, 423-441.
- [11] Neill, J. (1988) Testing for lack of fit in non-linear regression, *The Annals of Statistics*, **16**, 2, 733-740.

- [12] Nummelin, E. et Tuominen, P. (1982) Geometric ergodicity of Harris-recurrent Markov chains with application to renewal theory. *Stoch. Proc. Appl.*, **12**, 187-202.
- [13] Shorack, G. et Wellner, J. (1986) *Empirical Processes with Applications to Statistics*. Wiley : New York.
- [14] Strassen, V. (1965) The existence of probability measures with given marginals, *Ann. of Math. Stat.*, **36**, 423-439.