TEST OF NO-EFFECT HYPOTHESIS

DEMBO GADIAGA AND ROSARIA IGNACCOLO

Abstract. On se donne une suite de vecteurs aléatoires $(X_1, Y_1), ..., (X_n, Y_n)$ définies sur le même espace probabilisé (Ω, A, P). Après avoir considéré l'estimation de la fonction de regression $r(x)$, nous étudions le test d'hypothèse nulle "$r(x) = cste$", c'est à dire que X n'a pas d'effet en moyenne sur Y, dans deux situations où les variables aléatoires (X_i, Y_i) sont indépendantes ou forment un processus stationnaire et $\alpha-$mélangeant. Des lois limites sous diverses alternatives sont obtenues ainsi que des conditions nécessaires et suffisantes de convergence du test. Des simulations sont indiquées.

1. Introduction

The nonparametric test for a regression function has been studied by several authors for processes both in discrete and continuous time. Several methods have been used, among others the kernel, the orthogonal series and the least squares method. J. D. Hart(1997), in his book "Nonparametric smoothing and lack of fit tests", examines the relation between a variable Y and a deterministic variable x, expressed as $Y = r(x) + \varepsilon$, where r denotes the regression function and ε the error term. In later times Hart and Aerts et *al.* *(2000)* have developed tests under different assumptions. In Lee and Hart(2000), they use the trigonometric functions to estimate r and also supply a method to choose the order of the Fourier coefficients.

In our work we have assumed the existence of a relationship expressed by the model $Y = r(X) + \varepsilon$ where X is a random variable, to test the no-effect hypothesis $H_0 : r(x) = c$, where c is a constant. We have hence considered a projection estimator of $r(x)$ and defined the statistics of our test. We then analyze the asymptotic behavior of the test for independent and correlated observations, obtaining in both cases the limit distribution and the necessary and sufficient conditions for the consistency of the test.

2. Definition of the test

We consider the sequences of r.v.s (X_i, Y_i) with $1 \leq i \leq n$, defined on (Ω, A, P) with values in a mesurable space $(E \times I\!R, B \times B_{I\!R}, \mu \otimes \lambda)$ where μ is a $\sigma-$finite measure on E and λ the Lebesque measure. We suppose that the (X_i, Y_i) are identically distributed, and also that $X_i \backsim \mu$ with density $f(x)$.

Key words and phrases. régression, test non paramétrique, processus mélangeant.

This research was mainly conducted while the authors visited the Laboratoire de Statistiques Théoriques et Appliquées (LSTA) of the University "Pierre et Marie Curie" in Paris. The authors thank all LSTA members for their kindness and, in particular, Prof. Denis Bosq for his helpful advices and valuable comments.

Received on November 11, 2003. Revised version received on December 17, 2004

We consider $r(x) = E(Y \mid X = x)$ in the regression model $Y = r(X) + \varepsilon$ and we want to test the hypothesis $H_0 = c$ where c is a constant. Without loss of generality, we pose $c = 0$ (as c is known) since it is possible to apply the test on the obsevations $Y_i - c$ considering the model $Y - c = r(X) + \varepsilon$. Moreover, if $r_0(x)$ is a specified function one can use the test for the hypothesis $H_0 : r(x) = c + r_0(x)$ running on the observations $Y_i - c - r_0(X_i)$.

We suppose that $r \in L^2(\mu)$, so that one can write $r(x) =_{j=0}^{\infty} b_j e_j(x)$ with the fourier coefficients $b_j = \langle r, e_j \rangle = r(x) e_j(x) d\mu(x)$; For a fixed positive integer k, let $\{e_0, e_1, ..., e_k\}$ be an orthonormal system with $e_0 \equiv 1$, in $L^2(\mu)$, which generates a subspace E_k, with $\dim(E_k) = k + 1$. The estimator of the regression function by projection on E_k is defined by (see Bosq and Lecoutre (1987))

$$r_n(X, Y; x) = r_n(x) =_{j=0}^{k} \hat{b}_{jn} e_j(x)$$

where $\hat{b}_{jn} = \frac{1}{n}{}_{i=1}^{n} Y_i e_j(X_i)$ is the unbiased estimator of the fourier coefficients b_j. We observe that for $j = 0$ one has $\hat{b}_{0n} = \frac{1}{n}{}_{i=1}^{n} Y_i = \bar{Y}_n$.

To test the hypothesis $H_0 : r(x) = 0$, we consider the distance $d(r_n, 0) = \|r_n - 0\|$ in $L^2(\mu)$ between the estimated regression function and the true one under H_0, where $\|.\|$ denotes the $L^2(\mu) - norm$. Now we consider the statistic $R_n = \sqrt{n} r_n$ and its norm in $L^2(\mu)$, $\|R_n\| = \sqrt{n}\|r_n\|$, obtaining

$$\|R_n\|^2 = n_{j=0}^{k} \hat{b}_{jn}^2.$$

So the test rejects H_0 for large values of $\|R_n\|^2$. We observe that $\|R_n\|^2$ depends only on the estimated Fourier coefficients and that the null hypothesis $H_0 : r(x) = 0$ for every x versus $H_1 : r(x) \neq 0$ is equivalent to the system $H_0 : b_j = 0 \ \forall j \geq 0$ versus $H_1 : \exists j \geq 0 : b_j \neq 0$. Indeed we shall consider the alternative hypothesis $H_1(k) : r(x) \neq 0$ with r such that it exists $j \in \{0, 1, ..., k\}$ for which $b_j \neq 0$. Subsequently, we will adopt the following notations:

For each j, let us define the centered real random variables

$$D_{ij} = Y_i e_j(X_i) - E[Y_i e_j(X_i)], 1 \leq i \leq n$$

and for

$$n^{-\frac{1}{2}} S_{nj} = n^{\frac{-1}{2}}{}_{i=1}^{n} D_{ij} = \sqrt{n}(\widehat{b}_{in} - b_j),$$

let us consider also the vectors

$$n^{-\frac{1}{2}} S_n = \begin{pmatrix} n^{-\frac{1}{2}} S_{n0} \\ ... \\ n^{-\frac{1}{2}} S_{nk} \end{pmatrix} = \sqrt{n} \left[\begin{pmatrix} \widehat{b}_{0n} \\ ... \\ \widehat{b}_{kn} \end{pmatrix} - \begin{pmatrix} b_0 \\ ... \\ b_k \end{pmatrix} \right] = \sqrt{n}[B_n - b]$$

and let $D_i = \begin{pmatrix} D_{i0} \\ ... \\ D_{ik} \end{pmatrix}$, for $1 \leq i \leq n$ and the linear combination $v_i =_{j=0}^{k} c_j D_{ij} = C^T D_i$ with $C = (c_0, ..., c_k) \in \mathbb{R}^{k+1}$. Note finally matrix $\sum$ indicating the matrix with the elements defined by:

$$\sigma_{ij} = E\left[(Y_i e_j(X_i) - b_j)(Y_i e_l(X_i) - b_l)\right] \ .(1)$$

3. Large sample behaviour

We give here some results; for details see Ignaccolo (2002), or Gadiaga (2003). We first deal with the

INDEPENDENT DATA: We suppose now that the observations $(X_1, Y_1), ..., (X_n, Y_n)$ are independent identically distributed (iid).

Theorem 1: *Under $H_0 : r = 0$ and if $E\left(Y_i^2 \mid X_i\right) = \gamma$, where $\gamma > 0$ is a constant, one has*

1) $\|R_n\|^2 \xrightarrow{d} \gamma Q$, *where* $Q \sim \chi^2_{k+1}$.

2) Moreover, if $E\left(|D_{ij}|^3\right) < \infty$ *(where* $D_{ij} = Y_i e_j(X_i) - E(Y_i e_j(X_i))$*) for every* j *then*

$$\triangle_n = \sup_{t \in I\!R} \left| P\left(\|R_n\|^2 \leq t\right) - P(\gamma Q \leq t) \right| \leq c_0 \frac{(k+1)^3}{\gamma^{\frac{3}{2}}} {}^{k}_{j=0} E\left(|D_{1j}|^3\right) n^{-1/2} \quad (2)$$

where c_0 is a constant.

Proof:

1) Under H_0, the coefficients of Fourier are equal to zero, i.e $b = 0$. By using the central limit theorem in the multivariate case to the sequence $(D_i)_{1 \leq i \leq n}$ where the D_i are of the same law and independent, we can write:

$$\sqrt{n} B_n = n^{-\frac{1}{2}}{}_{i=1}^{n} D_i \xrightarrow{k} Z \sim N_{k+1}(0, \sum)$$

where $\sum$ denote the general element:

$$\sigma_{jl} = E[D_{ij} D_{il}] = E[(Y_i e_i(X_i))(Y_i e_l(X_i))]$$

and by using the conservation of the convergence in law by continuity, one has:

$$n B_n^T I_{k+1} B_n = n_{j=0}^{k} \widehat{b}^2_{jn} = \|R_n\|^2$$

i.e

$$\|R_n\|^2 = n B_n^T I_{k+1} B_n \xrightarrow{k} Z^T I_{k+1} Z$$

where I_{k+1} denote the identity matrix of order $k+1$. Moreover, the formula (1) becomes (under H_0)

$$\sigma_{ij} = E[Y_i^2 e_j(X_i) e_l(X_i)]$$

$$= E[E(Y_i^2 / X_i) e_j(X_i) e_l(X_i)]$$

and, with the condition $E(Y_i^2 / X_i) = \gamma$, one has:

$$\sigma_{il} = \begin{cases} 0 \text{ if } \ j \neq l \\ cE[e_i^2(X_i)] = \gamma \text{ if } \ j = l \end{cases}$$

i.e $\sum = \gamma I_{k+1}$. Then $Z^T I_{k+1} Z =_{j=0}^{k} \lambda_j U_j^2 = \gamma Q$ since $\lambda_j = \gamma$ are the eigenvalues of $\sum$ and the random variables $U_i \sim N(0, 1)$ are independent. consequently,

$$\|R_n\|^2 \xrightarrow{d} \gamma_{j=0}^{k} U_j^2 = \gamma Q.$$

2) By applyng Sazonov theorem (see appendix A) to the random variables D_i, with values in $I\!R^{k+1}$, i.i.d, centered, with the same matrix of variance-covariance as Z.

If we take $(t_0, t_1, ..., t_n)$ as the basic of unit eigenvectors of $\sum$, then it coincide with the canonical basic of $I\!R^{k+1}$; the scalar product $\langle D_{i,} t_j \rangle$ are the real random variables not correlated D_{ij} and the eigenvalues λ_j are equal to γ.

$$E[\langle D_1, t_j\rangle \langle D_1, t_l\rangle] = t_j^T \sum t_l = t_j^T (\lambda_l t_l) = \lambda_l \sigma_{jl} = \gamma \sigma_{jl},$$

then (see appendix A)

$$\rho_j^{(k)} = \frac{E\left[|\langle D_1, t_j\rangle|^3\right]}{\left[E\left[|\langle D_1, t_J\rangle|^2\right]\right]^{\frac{3}{2}}} = \frac{E\left[|D_{1j}|^3\right]}{\gamma^{\frac{3}{2}}}$$

Consequently, one has (2).

Remark:

We also prove (Th. 5.2.5) in Ignaccolo (2002) that under $H_1(k)$ one has $\|R_n\|^2 \xrightarrow{a.s} \infty$. So we can say that with $\alpha \in]0,1[$ such that

$$P\left(\gamma \chi_{k+1}^2 > w\right) = \alpha$$

i*t is possible to construct a test with rejection region* $\left\{(x_1, ..., x_n) : \|R_n\|^2 > w\right\}$ *with asymptotic size α and, under the null hypothesis, one has*

$$\left|P\left({}_{j=0}^{k}\hat{b}_{jn}^2 > \frac{w}{n}\right) - \alpha\right| = O\left(n^{-1/2}\right).$$

Since γ is unknown, to apply the test one can estimate it by

$$\hat{\gamma} = \frac{1}{k+1} \overset{k}{\underset{j=0}{}} \hat{\gamma}_j = \frac{1}{n(k+1)} \overset{k}{\underset{j=0}{}} {}_{i=1}^{n} Y_i^2 e_j^2(X_i) ..$$

Moreover (Th. 5.2.6 in Ignaccolo (2002)) *the test with rejection region* $\|R_n\|^2 > w_n$ *is consistent with respect to $H_1(k)$ if and only if* $w_n \underset{n\to\infty}{\to} \infty$ *and* $\frac{w_n}{n} \underset{n\to\infty}{\to} 0.$

Now we handle the **CORRELATED DATA** to *mod*el a weak dependence. In the following we suppose to have observations generated by $(X_t, Y_t)_{t\in Z}$ that is $\alpha - mixing$ with coefficients α_{XY}. We recall that a process $(X_t)_{t\in Z}$ is said strong mixing (or $\alpha - mixing$) if

$$\alpha(n) = \sup_{t\in Z} \sup_{A\in F_{-\infty}^t, B\in F_{t+n}^\infty} |P(A\cap B) - P(A)P(B)| \underset{n\to\infty}{\to} 0.$$

where $\alpha(n)$ are the mixing coefficients and $F_{-\infty}^t = \sigma(X_i, i \le t)$ denotes the $\sigma-$algebra generated by $(X_i, i \le t)$ and $F_{t+n}^\infty = \sigma(X_i, i \ge t+n)$. Now we pose $D_{ij} = Y_i e_j(X_i) - E(Y_i e_j(X_i))$, for $1 \le i \le n$, and $V_i = {}_{j=0}^{k} c_j D_{ij} = c^T D_i$ with $c = (c_0, ..., c_k)^T \in I\!R^{k+1}$ and we consider the assumptions:

(A) $E\left(|D_{1j}|^{2+\delta_D}\right) < \infty$ for $\delta_D > 0$;

(B) $E\left(|V_1|^{2+\delta_V}\right) < \infty$ for $\delta_V > 0$;

(C) $\sigma^2 > 0$ where $\sigma^2 = E\left(|V_1|^2\right) + 2_{i=2}^{\infty} E\left(V_1 V_i\right) =_{i=-\infty}^{+\infty} E\left(V_1 V_i\right)$.

Theorem 2. *Under $H_0 : r(x) = 0$ and under the assumptions (B)-(C), we set*

$$\triangle_n^* = \sup_{u \in I\!R} \left| P\left(\|R_n\|^2 \le u\right) - P\left(\|U^*\|^2 \le u\right) \right|$$

with $U^* =_{j=0}^{k} \lambda_j^* U_j e_j$, where λ_j^{*2} are the eigenvalues of the matrix * defined by

$$\sigma_{jl}^* = E\left(Y_1^2 e_j\left(X_1\right) e_l\left(X_1\right)\right) +_{i=2}^{\infty} E\left(Y_1 e_j\left(X_1\right) Y_i e_l\left(X_1\right)\right) +_{i=2}^{\infty} E\left(Y_i e_j\left(X_i\right) Y_1 e_l\left(X_1\right)\right)$$

and the r.v.s $U_j \sim N(0,1)$ are independent. Then

i) if $\alpha_{XY}(n) = O\left(n^{-\beta(2+\delta)(1+\delta)/\delta^2}\right)$ for some $\beta > 1$, then there exists a constant γ_1 such that

$$\triangle_n^* \le \gamma_1 n^{-\frac{\delta(\beta-1)}{2(\beta+1)}},$$

ii) if $\alpha_{XY}(n) = O\left(e^{-\beta n}\right)$ for some $\beta > 1$, then there exists a constant γ_2 such that

$$\triangle_n^* \le \gamma_2 n^{-\frac{\delta}{2}} \log^{1+\delta} n,$$

with $\delta = \max\left(\delta_Y, \delta_V\right)$. Moreover,

$$\|R_n\|^2 \xrightarrow{d}{}^{k}_{j=0} \lambda_j^{*2} U_j^2.$$

Proof:

We will examine only case i), case ii) shows ourselves in a similar way. With the notations and the conditions above, one can apply Tikhomirov theorem (appendix B) to the sequence V_i, which is centered and real values by definition, stationary and $\alpha - mixing$ with $\alpha_V(n) = \alpha_X(n)$, since we have $V_i = C^T X_i = f(X_i)$ with f mesurable; then there exists a constant γ_1 such that

$$\sup_{t \in I\!R} \left| P\left[n^{-\frac{1}{2}}{}_{i=1}^{n} V_i \le t\right] - P[N_\sigma \le t] \right| \le \gamma_1 n - \frac{\delta_v(\beta_v - 1)}{2(\beta_v + 1)}$$

where $N_\sigma \sim N(0, \sigma^2)$ since $\sigma^2 < +\infty$; but,

$${}_{i=1}^{n} V_i = c_1{}_{i=1}^{n} X_{i1} + \ldots + c_n{}_{i=1}^{n} X_{ik}$$

$$=_{i=1}^{n} c_i \delta_{nj} = C^T \delta_n.$$

Then, $\forall C \in I\!R^k$, $C^T \delta_n \xrightarrow{d} N(0, \sigma^2)$ and if $\sigma^2 = C^T \sum C$ where $\sum$ is the matrix of variance-covariance of vector Z which is definite positive since $\sigma^2 > 0$, by condition (C), on has:

$$C^T \delta_n \xrightarrow{d} N\left(0, C^T \sum C\right) \text{ and } C^T \delta_n \xrightarrow{d} C^T Z. \qquad (3)$$

It remains to be seen that $\sigma^2 = C^T \sum C$ with $\sum$ symmetrical, one has:

$$\sigma^2 = E(V_1^2) + 2_{i=2}^{\infty} E[V_1 V_i]$$

but,

$$E[V_1 V_i] = E\left[[C^T X_1][C^T X_i]^T\right] = E\left[[C^T X_1 X_i{}^T C\right] = C^T E\left[X_1 X_i{}^T\right] C = C^T \sum_{1,i} C.$$

In the same way, $E[V_i V_1] = C^T \sum_{i,1} C$ where $\sum_{1,i} = E[X_1 X_i^T]$. Then $\sum_{i,1} = E[X_i X_1^T]$. Further

$$\sigma^2 = C^T \sum_{1,1} C +_{i=2}^{\infty} C^T \sum_{1,i} C +_{i=2}^{\infty} C^T \sum_{i,1} C$$

$$= C^T [\sum_{1,1} +_{i=2}^{\infty} \sum_{1,i} +_{i=2}^{\infty} \sum_{i,1}] C = C^T \sum C.$$

By (3), by applying the criterion of Cramer Wold, on has:

$$n^{-\frac{1}{2}} S_n \xrightarrow{d} Z = N_k(0, \sum).$$

Under H_0, the coefficients of Fourier are equal to zero, i.e $b = 0$, then $D_{ij} = Y_i c_j(X_i)$ and consequently,

$$\triangle_n^* \leq \gamma_1 n^{-\dfrac{\delta(\beta-1)}{2(\beta+1)}}$$

where $\|R_n\|^2 = n_{j=0}^{n} \widehat{b}_{jn}^2 \xrightarrow{d}{}_{j=0}^{k} \lambda_j^{*2} U_j^2$.

Remark:

We prove again (Th. 5.3.3 in Ignaccolo (2002)) that under $H_1(k)$ one has $\|R_n\|^2 \xrightarrow{a.s} \infty$. Hence if we get $\alpha \in]0,1[$ such that
$P\left({}_{j=0}^{k} \lambda_j^{*2} U_j^2 > w\right) = \alpha.$

We can construct a test with rejection region $\left\{(x_1, ..., x_n) : \|R_n\|^2 > w\right\}$ with asymptotic size α. Raeally the estimation of the eigenvalues λ_j^{*2} needs. Then we

define an estimator of σ_{jl}^* as $\hat{\sigma}_{jl} ={}_{v=-l_n}^{l_n} w_n(v)\, \hat{\sigma}_{jl}(v)$, where $\hat{\sigma}_{jl}(v)$ is the classic estimator of the crossed covariance of the stationary bivariate process $(D_{tj}, D_{tl})_{t\in Z}$, $w_n(v)$ are weights satisfying some conditions and l_n is a sequence of positive integers such that $l_n < n$ and $\lim_{n\to\infty} l_n = \lim_{n\to\infty} \frac{n}{l_n}$. Under suitable conditions $\hat{\sigma}_{jl}$ is consistent, so we can use the estimated eigenvalues to have an approximation of the limit distribution (for further details see Ignaccolo (2002)).

To obtain the consistency of the test we set a boundedness condition, that is:

$$(\mathbf{D}) \quad \sup_{1\leq i\leq n} \sup_{x\in E} \max_{0\leq j\leq k} |Y_i e_j(x)| = M < \infty,$$

and we have the following result :

Theorem 3. *Under the assumptions (B)-(D) and if* ${}_{i=1}^{\infty}\alpha^*(i) < \infty$, *where* $\alpha^*(i)$ *are the mixing coefficients for the r.v.s* $Y_i e_j(X_i)$, *the test with rejection region* $\|R_n\|^2 > w_n$ *is consistent with respect to* $H_1(k)$ *if and only if* $w_n \underset{n\to\infty}{\to} \infty$ *and* $\dfrac{w_n}{n} \underset{n\to\infty}{\to} 0$.

Proof: The application of the theorem is equivalent to the following points:
1) $\alpha_n \to 0 \Longleftrightarrow w_n \to +\infty$
2) ${}_{i=1}^{\infty}\alpha^*(i) < \infty$ and $\dfrac{w_n}{n} \underset{n\to\infty}{\to} 0 \Longrightarrow \beta_n(r_1) \to 1$
3) $\beta_n(r_1) \to 1 \Rightarrow \dfrac{w_n}{n} \underset{n\to\infty}{\to} 0$
where $\beta_n(r_1)$ denote the power of the test and α_n is level:

$$\alpha_n = P_{r=0}[\|R_n\|^2 > w_n]$$

$$\beta_n(r_1) = P_{r\neq 0}[\|R_n\|^2 > w_n] :$$

It is thus enough to prove the 3 points above.
1) For the random variable $\|U^*\|^2 = {}_{j=0}^{k} \lambda_j^{*2} U_j^2$, one has:

$$P[\|U^*\|^2 > w_n] \leq \frac{{}_{j=0}^{k}\lambda_j^{*2}}{w_n} \to 0 \Longleftrightarrow w_n \to +\infty.$$

According to item i) of theorem 2, one has:

$$\left|\alpha_n - P[\|U^*\|^2 > w_n]\right| \leq \gamma_1 n^{-\frac{\delta(\beta-1)}{2(\beta+1)}}$$

then $\alpha_n \to 0$ if and only if $w_n \to \infty$. What we get by considering also item ii) of theorem 2.

2) According to $H_1(k)$, there is $j \in \{0, 1, ..., k\}$, that is to say j_0 such that $b_{j_0} = E[Ye_{j_0}(X)] \neq 0$.

Let $m > 1$ be an integer and $N(b_{j_0}, m)$ the smollest integer such that $\dfrac{w_n}{n} \leq \dfrac{b_{j_0}^2}{m^2}$ for $n \geq N(b_{j_0}, m)$. As

$$\|R_n\|^2 \leq w_n \Rightarrow_{j=0}^{k} \widehat{b}_{jn}^2 \leq \frac{w_n}{n} \Rightarrow \left|\widehat{b}_{j_0 n}\right| \leq \sqrt{\frac{w_n}{n}}$$

for $n \geq N(b_{j_0}, m)$, one has

$$\left|\widehat{b}_{j_0 n}\right| \leq \frac{|b_{j_0 n}|}{m};$$

from where,

$$\left|\widehat{b}_{j_0} - b_{j_0}\right| \geq |b_{j_0}| - \left|\widehat{b}_{j_0}\right| \geq |b_{j_0}| - \frac{|b_{j_0}|}{m} = |b_{j_0}| \frac{m-1}{m}.$$

Consequently, for $n \geq N(b_{j_0}, m)$,

Dembo Gadiaga & Rosario Ignaccolo, Afrika Statistika, Vol.1, n°1, 2005, pp.67-76
Test of No-effect Hypothesis

$$\begin{aligned} 1-\beta_n(r_1) &= P[\|R_n\|^2 \leq w_n] \leq P[\left|\widehat{b}_{j_0n} - b_{j_0}\right| \geq |b_{j_0}| \frac{m-1}{m}] \\ &\leq \frac{E[\left|\widehat{b}_{j_0n} - b_{j_0}\right|^2]}{|b_{j_0}|^2 (\frac{m-1}{m})^2} = \frac{Var(\widehat{b}_{j_0n})}{b_{j_0}^2} \frac{m^2}{(m-1)^2} \end{aligned}$$

by using Bienaymé-Tchebychev inequality. Let us raise $Var(\widehat{b}_{j_0n})$; by posing $R(i) = Cov(Y_1 e_{j_0}(X_1), Y_{i+1} e_{j_0}(X_{i+1}))$,

$$\begin{aligned} Var(\widehat{b}_{j_0n}) &= \frac{1}{n^2} Var(_{i=1}^{n} Y_i e_{j_0}(X_i)) \\ &= \frac{1}{n}{}_{|i|\leq n-1}(1-\frac{|i|}{n})R(i) \\ &\leq \frac{1}{n}[|R(0)| + 2_{i=1}^{n-1}|R(i)|] \end{aligned}$$

$$Var(\widehat{b}_{j_0n}) \leq \frac{1}{n}[|R(0)| + 8M^2_{i=1}{}^{\infty}|\alpha^*(i)|]$$

where in the last majoration, we have used exponential inequality:

$$|Cov(Y_s, Y_{s-t})| \leq 4\|Y_s\|_\infty \|Y_{s+t}\|_\infty \alpha(t)$$

to the random variables $Y_i e_j(X_i)$ bounded by assumptions.

Then, $1-\beta_n(r_1) \to 0$ for $n \to +\infty$.

3) One reasons by adjonction by supposing that $\frac{w_n}{n} \nrightarrow 0$; then, there is $\eta > 0$ and an infinite part $N_1 \subset I\!N$ such that $\frac{w_n}{n} > \eta$ for $n \in N_1$. Taking $\eta \in H_1(k)$ such that $0 <_{j=0}^{k} b_j^2 < \eta$, then existence is done because for a given $\eta > 0$, take $r_i(x) = b_1 e_1(x)$ in $H_1(k)$, choosing $0 < b_1 < \sqrt{\eta}$, then $_{j=0}^{k} b_j^2 < \eta$. The condition $\beta_n(r_1) \to 1$ give for $n \in N_1$,

$$P[_{j=0}^{k}\widehat{b}_{jn}^2 \leq \eta] \leq P[_{j=0}^{k}\widehat{b}_{jn}^2 \leq \frac{w_n}{n}] \underset{n\to+\infty}{\to 0}.$$

In addition

$$_{j=0}^{k}\widehat{b}_{jn}^2 \overset{a.s}{\to}{}_{j=0}^{k} b_j^2 < \eta$$

from where,

$$P[_{j=0}^{k}\widehat{b}_{jn}^2 \leq \eta] \leq P[_{j=0}^{k}(\widehat{b}_{jn}^2 - b_j^2) < \eta -_{j=0}^{k} b_j^2] \underset{n\to+\infty}{\to 1}.$$

which is absurd.

4. Conclusion

This work lies in the wake of the functional tests of fit associated with the projection estimators, introduced by Bosq (for general results see Bosq (2002)). The results of this work bring us the limit distibutions for the case of no effect hypothesis, for correlated data, the method requires an estimation of the eigenvalues and the deermination of the quantiles of a linear combination of χ_1^2. Simulations to evaluate the empirical power of the test are in progress.

5. Appendix

A- Theorem of Sazonov (Sazonov, 1968) [9,10]

Let $(X_n)_{n\geq 1}$ be a sequence of random variable of values in $\mathbb{R}^k$, independent and of the same law ν, centered and admitting one moment of order 3. Let C be the class of convex measurable of $\mathbb{R}^k$; and let $t = (t_1, t_2, ..., t_k)$ be a finite part of $\mathbb{R}^k$ such that the scalar productI $(\langle X_1, t_j\rangle), j = 1, , ..., k$ are not correlated random variables. Then

$$\sup_{B\in C} |P_n(B) - N(B)| \leq c_0 k^3_{j=1}{}^k \rho_j^{(t)} n^{-\frac{1}{2}}, \quad n \geq 1$$

where P_n denote the law of $\frac{1}{\sqrt{n}}{}_{i=1}^{k} X_i$, N the normal law having the same moments of order 1 and 2 like ν and where we posed:

$$\rho_j^{(t)} = \frac{E[|\langle X_1, t_j\rangle|^3]}{(E[|\langle X_1, t_j\rangle|^2])^{\frac{3}{2}}}$$

and where c_0 denote a universal constant.

B- Theorem of Tikhomirov

Let $(X_t)_{t\in\mathbb{R}}$ a process centered, stationary, $\alpha - mixing$ and of real values. Let us pose

$$\Delta_n = \sup_{t\in\mathbb{R}} \left| P(n^{-\frac{1}{2}} S_n \leq t) - P(N_\sigma \leq t) \right|$$

where $N_\sigma \sim N(0, \sigma^2)$, $S_n = \frac{1}{\sqrt{n}}{}_{i=1}^{n} X_i$.

If $E[|X_1|^{2+\delta}] < +\infty$ for $\delta > 0$, and if $\sigma^2 > 0$,one has:

i) If $\alpha(n) = O(n^{-\beta(2+\delta)(1+\delta)/\delta^2})$ for $\beta > 1$ then there is a constant γ_1 such that

$$\Delta_n \leq \gamma_1 n^{-\frac{\delta(\beta-1)}{2(\beta+1)}}.$$

ii) If $\alpha(n) = O(e^{-\beta n})$ for $\beta > 1$, then there is a constant γ_2 such that $\Delta_n \leq \gamma_2 n^{-\frac{\delta}{2}} \log^{1+\delta} n$.

Acknowledgment: We thank the referee we allowed us through is observations, for improving the drafting of the paper.

References

[1] Aerts M., Claeskens G. and Hart J.D. (2000) Testing lack of fit in multiple regression, Biometrika, 87, 2,405-424.
[2] Bosq D. (2002) Functional test of fit, in: Goodness-of-Fit Tests and Model Validity, Birkhaüser, Boston, chapter 25, 341-356, in: C. Huber-Carol et al.: (Eds.).
[3] Bosq D. and Lecoutre J.P. (1987) Théorie de l'estimation fonctionnelle, Economica, Paris.
[4] Gadiaga D. (2003) Tests fonctionnels d'ajustement et de non influence pour des variables aléatoires dépendantes, thèse d'Etat Université de Ouagadougou. Nov 2003.
[5] Hart J.D. (1997) Nonparametric smothing and lack-of-fit tests, Springer-Verlag, New York.
[6] Ignaccolo R. (2002) Tests d'ajustement fonctionnels pour des observations correlées, Ph.D. thesis, Università degli Studi di Milano and Université Paris VI. Nov 2002.
[7] Ignaccolo R. (2003) Goodness of fit tests for dependent data, journal of nonparametric Statistics, in corso di stampa, 1-20.
[8] Lee G. and Hart J.D. (2000) Model selection criteria with data dependent penalty, with applications to data-driven Neyman smooth tests, J. Nonparametr. Statist., 12, 5, 683-707.
[9] Sazonov V.V, On w^2 criterion, Sankhya $A - 30$ (1968),$201 - 210$.
[10] Sazonov V.V, On the multi-dimensional central limit theorem, Sankhya A 30 (1968), 191-204.
[11] Thikomirov A.N, Convergence rate in the central limit theoem for weakly dependent random variables, theor. veroyaknosk. In primenen. 25 (1980), n$^0 4, 800 - 818$.

UFR SEG, Université de Ouagadougou, Burkina Faso.
E-mail address: gdembo@hotmail.com

Current address: Dipartimento di Statistica e Matematica Applicata, Università degli Studi di Torino, Italia.
E-mail address: Ignaccolo@econ.unito.it