

ANNALES DE L'I. H. P., SECTION B

JOSÉ RAFAEL LEON R.

Asymptotic behaviour of the quadratic measure of deviation of multivariate density estimates

Annales de l'I. H. P., section B, tome 19, n° 3 (1983), p. 297-309

http://www.numdam.org/item?id=AIHPB_1983__19_3_297_0

© Gauthier-Villars, 1983, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P., section B » (<http://www.elsevier.com/locate/anihpb>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

Asymptotic behaviour of the quadratic measure of deviation of multivariate density estimates

by

José Rafael LEON R.

Departamento de Matemáticas, Facultad de Ciencias,
Universidad Central de Venezuela, Caracas, Aptdo. n° 21201

RÉSUMÉ. — Nous obtenons un test d'adéquation de la distribution asymptotique de $\|\hat{f}_{n,N} - f_n\|^2$ et nous prouvons également que la statistique considérée est asymptotiquement gaussienne sous les hypothèses de contiguïté de la forme $f_N = f^0 + \delta_N \phi$, $\phi \in L^2(\mu)$, $\delta_N \downarrow 0$.

ABSTRACT. — We obtain a test of goodness of fit from the asymptotic distribution of $\|\hat{f}_{n,N} - f_n\|^2$ and we also prove that the statistic under consideration is asymptotically gaussian under contiguous alternatives of the form $f_N = f^0 + \delta_N \phi$, $\phi \in L^2(\mu)$, $\delta_N \downarrow 0$.

Key words and phrases: Multidimensional density estimates, quadratic measure, asymptotic distribution, test of goodness of fit.

1. INTRODUCTION

Set $X_1, X_2, \dots, X_N, \dots$ be a sequence of independent identically distributed random vector with values in \mathbb{R}^p . We shall suppose that their common distribution has a density f with respect to Lebesgue measure and that $f \in L^2(\mu)$, where μ is a Borel probability measure on \mathbb{R}^p with density $r(x)$ with respect to Lebesgue measure.

If $\{\phi_j\}_{j=1}^{\infty}$ is complete orthonormal system in $L^2(\mu)$, the n -th partial sum of the respective Fourier series for f is

$$f_n(x) = \sum_{j=1}^n a_j \phi_j(x) \quad x \in \mathbb{R}^p$$

where

$$a_j = \int_{\mathbb{R}^p} f(x) \phi_j(x) r(x) dx = E \alpha_j(X_1) \quad \alpha_j(x) = \phi_j(x) r(x)$$

$\hat{\text{Cencov}}$ [2] defines the following estimator of f_n :

$$\hat{f}_{n,N}(x) = \sum_{j=1}^n \hat{a}_j \phi_j(x)$$

where the \hat{a}_j 's are estimators of a_j defined by

$$\hat{a}_j = \int_{\mathbb{R}^p} \alpha_j(x) dF_N(x)$$

F_N being, as usual, the empirical distribution function of the sample X_1, X_2, \dots, X_N .

The aim of this paper is to give conditions under which

$$\frac{N}{n^{1/2}} \|\hat{f}_{n,N} - f_n\|^2.$$

When appropriately centered, has gaussian asymptotic distribution. The method we use is inspired by Naradaya [8], although instead of using the strong approximation of the empirical process by a Brownian bridge, we approximate the estimator by functions of Gaussian variables with values in $L^2(\mu)$. For these ones, the result follows from the central limit theorem on the real line, and the approximation allows to study the

behaviour of $\frac{N}{\sqrt{n}} \|\hat{f}_{n,N} - f_n\|^2$.

The result is applied to various complete orthonormal sets.

Finally, we consider tests of goodness of fit upon $\|\hat{f}_{n,N} - f_n\|^2$ and the behaviour of $g(n) = \|f_n - f\|^2$, which permit together to study the asymptotic behaviour of the statistic $\|\hat{f}_{n,N} - f\|^2$.

2. ASSOCIATED GAUSSIAN VARIABLES

We have defined

$$\hat{f}_{n,N}(x) = \sum_{j=1}^n \hat{a}_j \phi_j(x)$$

If we put

$$Y_{n,k}(x) = \sum_{j=1}^n (\alpha_j(X_k) - a_j) \phi_j(x)$$

it is clear that

$$\sqrt{N}(\hat{f}_{n,N} - f_n) = \frac{1}{\sqrt{N}} \sum_{k=1}^N Y_{n,k}(x)$$

If we consider the $Y_{n,k}$ as independent identically distributed random variables with values in $L^2(\mu)$ we have

$$E(Y_{n,k}) = 0$$

and

$$\Gamma_n(g, h) = E \{ (g, Y_{n,k})(h, Y_{n,k}) \}$$

where Γ_n is the covariance of $Y_{n,k}$ and (\cdot, \cdot) is the scalar product in $L^2(\mu)$.

It follows that

$$\Gamma_n(\phi_i, \phi_j) = \int_{\mathbb{R}^p} f(x) \phi_i(x) \phi_j(x) r^2(x) dx - a_i a_j \quad i, j = 1, 2, \dots, n.$$

Define the centered gaussian random variable $Z_{1,n}$ with values in $L^2(\mu)$ by

$$Z_{1,n} = \sum_{i=1}^n \xi_i \phi_i$$

in such a way that $Z_{1,n}$ and $Y_{n,k}$ have the same covariance. Now, if ξ_0 is a normalized gaussian real random variable, independent from $\{ \xi_i \}_{i=1}^n$, consider the random variable (with values in $L^2(\mu)$):

$$Z_{2,n} = Z_{1,n} + \xi_0 \sum_{i=1}^n a_i \phi_i.$$

Then $Z_{2,n}$ is gaussian, $E(Z_{2,n}) = 0$ and its covariance $\Gamma_n^{(2)}$ satisfies

$$\Gamma_n^{(2)}(\phi_i, \phi_j) = \int_{\mathbb{R}^p} \phi_i(x)\phi_j(x)r^2(x)f(x)dx$$

LEMMA 2.1. — If $\|fr\|_\infty < \infty$, then

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} E | \|Z_{2,n}\|^2 - \|Z_{1,n}\|^2 | = 0$$

Proof. — $E | \|Z_{2,n}\|^2 - \|Z_{1,n}\|^2 | \leq \sum_{j=1}^n a_j^2 + 2E |\xi_0| E | \sum_{j=1}^n \xi_j a_j |$
and since $E(|\xi_0|) < 1$,

$$\left(E \left| \sum_{j=1}^n \xi_j a_j \right| \right)^2 \leq \int_{\mathbb{R}^p} \left(\sum_{i=1}^n \phi_i(x) a_i \right)^2 r^2(x) f(x) dx \leq \|fr\|_\infty \|f\|^2.$$

The result now follows from

$$E | \|Z_{2,n}\|^2 - \|Z_{1,n}\|^2 | \leq \|f\| (\|f\| + 2 \|fr\|_\infty^{1/2}).$$

Before studying the asymptotic behaviour of $Z_{1,n}$ let us define:

- i) $A_n = (\Gamma_n^{(2)}(\phi_i, \phi_j)) = (C_{ij})$
- ii) $\Delta_n = \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}^p} \alpha_i^2(x) f(x) dx = \frac{1}{n} \text{tr} (A_n)$
- iii) $S_m(n) = \sum_{i_1=1}^n \dots \sum_{i_m=1}^n C_{i_1 i_2} C_{i_2 i_3} \dots C_{i_m i_1}$, evidently $S_m(n) = \text{Tr} (A_n^m)$.
- iv) $\sigma_n^2 = \frac{2}{n} \sum_{i=1}^n \sum_{j=1}^n \left(\int_{\mathbb{R}^p} \phi_i(x)\phi_j(x)r^2(x)f(x)dx \right)^2$

Note that

$$\sigma_n^2 = \frac{2}{n} S_2(n) = \frac{2}{n} \sum_{i=1}^n \lambda_{i,n}^2$$

if $\{\lambda_{i,n}\}_{i=1}^n$ are the eigenvalues of A_n .

LEMMA 2.2. — Suppose that there exists $m \geq 3$ such that $\frac{1}{n} S_m = 0(1)$. Then, if $\lim_{n \rightarrow \infty} \sigma_n^2 = \sigma^2 > 0$, we have

$$\lim_{n \rightarrow \infty} \frac{n\sigma_n^2}{(\max_{1 \leq i \leq n} \lambda_{i,n})^2} = \infty$$

Proof. —
$$\left(\sup_{1 \leq i \leq n} \lambda_{i,n} \right)^m \leq \sum_{i=1}^m \lambda_{i,n}^m = S_m(n)$$

So that

$$\frac{n\sigma_n^2}{\left(\sup_{1 \leq i \leq n} \lambda_{i,n} \right)^2} \geq \frac{n^{\left(\frac{m-2}{m}\right)} \sigma_n^2}{\left(\frac{1}{n} S_m(n)\right)^{2/m}} .$$

The result follows letting $n \rightarrow \infty$.

THEOREM 2.3. — Under the same hypothesis of Lemma 2.2 we have

$$W - \lim_{n \rightarrow \infty} \frac{1}{n^{1/2}} (\|Z_{2,n}\|^2 - E \|Z_{2,n}\|^2) = N(0, \sigma^2).$$

Proof. — Since $Z_{2,n}$ is gaussian we can find a basis $\{e_1, \dots, e_n\}$ such that

$$Z_{2,n} = \lambda_{1,n}^{1/2} \gamma_1 e_1 + \dots + \lambda_{n,n}^{1/2} \gamma_n e_n .$$

where $\gamma_1, \dots, \gamma_n$ are normalized independent gaussian random variables. So

$$\|Z_{2,n}\|^2 = \lambda_{1,n} \gamma_1^2 + \dots + \lambda_{n,n} \gamma_n^2$$

and

$$n^{-1/2} (\|Z_{2,n}\|^2 - E \|Z_{2,n}\|^2) = n^{-1/2} \sum_{i=1}^n \lambda_{i,n} (\gamma_i^2 - 1).$$

The result follows from

$$\lim_{n \rightarrow \infty} \frac{2}{n} \sum_{i=1}^n \lambda_{i,n}^2 = \sigma^2 .$$

Together with Lemma 2.2 and Lindeberg's Theorem on the line.

REMARK. — From Lemma 2.1 and $n^{-1/2} E \|Z_{2,n}\|^2 = \sqrt{n} \Delta_n$ we obtain

$$W - \lim_{n \rightarrow \infty} (n^{-1/2} \|Z_{1,n}\|^2 - n^{1/2} \Delta_n) = N(0, \sigma^2).$$

3. MAIN THEOREM

The following Theorem is due to Kuelbs and Kurtz [7] see also Giné and León [4]. The statement is adapted to our present needs.

THEOREM 3.1. — Let $\{Y_i\}_{i=1}^n$ be independent identically distributed

random variables with values in $L^2(\mu)$, $E(Y_1) = 0$, $E(\|Y_1\|^3) < \infty$ and Z_1 a centered gaussian variable with the same covariance as Y_1 . Then, for each t and $\delta > 0$,

$$\left| \mathbf{P} \left\{ \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^N Y_i \right\| \leq t \right\} - \mathbf{P} \{ \|Z_1\| < t \} \right| \\ = 0 \left(\delta^{-3} \frac{E \|Y_1\|^3}{\sqrt{N}} \right) + \mathbf{P} \{ \| \|Z_1\| - t \} \leq \delta \}$$

holds true.

We now prove our main result:

THEOREM 3.2. — Suppose that for some $\alpha > 0$

$$E(\|Y_{n,1}\|^3) = o(N^{1/2}) \quad \text{if} \quad n = 0(N^\alpha)$$

If, additionally, $\|f_r\|_r < \infty$, $\sigma_n^2 \rightarrow \sigma^2 > 0$ and $\frac{1}{n} S_m(n) = 0(1)$ for some $m \geq 3$, then

$$W - \lim_{n \rightarrow \infty} \left[\frac{N}{n^{1/2}} \| \hat{f}_{n,N} - f_n \|^2 - n^{1/2} \Delta_n \right] = N(0, \sigma^2)$$

Proof. — Define $G_{n,N}(t)$ and $G_n(t)$ in the following way:

$$\begin{aligned} G_{n,N}(t) &= \mathbf{P} \left\{ \frac{N}{n^{1/2}} \| \hat{f}_{n,N} - f_n \|^2 - n^{1/2} \Delta_n \leq t \right\} \\ &= \mathbf{P} \left\{ \left\| \frac{1}{\sqrt{N}} \sum_{k=1}^n \frac{Y_{n,k}}{n^{1/4}} \right\| \leq (t + n^{1/2} \Delta_n)^{1/2} \right\} \\ G_n(t) &= \mathbf{P} \left\{ \frac{1}{n^{1/2}} \| Z_{1,n} \|^2 - n^{1/2} \Delta_n \leq t \right\} \\ &= \mathbf{P} \left\{ \left\| \frac{1}{n^{1/4}} Z_{1,n} \right\| \leq (t + n^{1/2} \Delta_n)^{1/2} \right\}. \end{aligned}$$

By theorem 3.1.

$$|G_{n,N}(t) - G_n(t)| \\ = 0 \left(\delta^{-3} \frac{E \|Y_{n,1}\|^3}{n^{3/4} N^{1/2}} \right) + \mathbf{P} \left\{ \left| \frac{1}{n^{1/4}} \| Z_{1,n} \| - (t + n^{1/2} \Delta_n)^{1/2} \right| \leq \delta \right\}.$$

But

$$\begin{aligned} \mathbf{P} \left\{ \left| \frac{1}{n^{1/4}} \| Z_{1,n} \| - (t + n^{1/2} \Delta_n)^{1/2} \right| \leq \delta \right\} &= \mathbf{P} \left\{ \delta^2 - 2\delta \sqrt{t + n^{1/2} \Delta_n} + t \right. \\ &\leq \frac{1}{n^{1/2}} \| Z_{1,n} \|^2 - n^{1/2} \Delta_n \leq \delta^2 + 2\delta \sqrt{t + n^{1/2} \Delta_n} + t \left. \right\}. \end{aligned}$$

If we choose $\delta = \delta(n)$ such that $\delta^2 n^{1/2} \rightarrow 0$, the remark following theorem 2.3 and the fact that Δ_n is bounded, since

$$\Delta_n \leq \left(\frac{1}{n} \sum_{i=1}^n \lambda_{i,n}^2 \right)^{1/2} \rightarrow \frac{1}{\sqrt{2}} \sigma,$$

imply:

$$P\left(\left| \frac{1}{n^{1/4}} \|Z_{1,n}\| - (t + n^{1/2} \Delta_n)^{1/2} \right| \leq \delta \right) \rightarrow \phi_\sigma(t) - \phi_\sigma(t) = 0.$$

Where ϕ_σ denotes the normal $(0, \sigma^2)$ distribution.

Finally, if we choose $\delta(n) = n^{-1/4} \theta_n$ with $\theta_n \rightarrow 0$, $\delta^2 n^{1/2} = \theta_n \rightarrow 0$ holds, and the theorem will follow if the first term in (3.1) tends to zero. This is achieved if θ_n tends to zero slowly enough.

COROLLARY 3.3. — If in addition to the hypothesis of theorem 3.2, one has $g(n) = \|f_n - f\|^2 = o(n^{1/2} N^{-1})$, then

$$W - \lim_{n \rightarrow \infty} \left[\frac{N}{n^{1/2}} \|\hat{f}_{n,N} - f\|^2 - n^{1/2} \Delta_n \right] = N(0, \sigma^2).$$

We shall use the following theorem to verify that

$$E(\|Y_{n,1}\|^3) = o(N^{1/2}).$$

THEOREM 3.4. — Define $v_i = \sup_x |\alpha_i(x)|$. Suppose that

$$\sum_{i=1}^n v_i^2 = o(n^{-2} N) \quad \text{and} \quad \|fr\|_\infty < \infty$$

then

$$E(\|Y_{n,1}\|^3) = o(N^{1/2})$$

Proof. —

$$E(\|Y_{n,1}\|^3) = E\left(\sum_{j=1}^n (\alpha_j(X_1) - a_j)^2 \right)^{3/2} \leq 2^{5/2} E\left(\sum_{j=1}^n \alpha_j^2(X_1) \right)^{3/2}$$

now

$$\sum_{j=1}^n \alpha_j^2(X_1) \leq \sum_{j=1}^n v_j^{2/3} \alpha_j^{4/3}(X_1)$$

on applying Hölder's inequality it follows that

$$E\left(\sum_{j=1}^n \alpha_j^2(X_1)\right)^{3/2} \leq \left(\sum_{j=1}^n v_j^2\right)^{1/2} E\left(\sum_{j=1}^n \alpha_j^2(X_1)\right)$$

but

$$E\left(\sum_{j=1}^n \alpha_j^2(X_1)\right) = o(n)$$

then

$$E(\|Y_{n,1}\|^3) \leq K\left(n^2 \sum_{j=1}^n v_j^2\right)^{1/2} = o(N^{1/2})$$

according to the hypothesis.

4. EXAMPLES AND APPLICATIONS

a) Let us consider the space $L^2((-\pi, \pi)^2)$ with respect to Lebesgue measure, and the complete orthonormal set $\left\{\frac{1}{2\pi} e^{i(mx+ny)}\right\}_{(m,n) \in \mathbb{Z}^2}$. With the obvious modifications in the previous sections to be able to include complex valued functions f , we may study the estimators

$$\hat{f}_{n,N}(x, y) = \sum_{j,k=-n}^n \hat{C}_{k,j} e^{i(kx+jy)}$$

(Here, $C_{k,j}$ are the Fourier coefficients of f , and $\hat{C}_{k,j}$ their estimators).

We easily verify

$$\Delta_n = \frac{2}{4\pi^2}$$

and

$$\sigma_n^2 = \frac{1}{2\pi^2(2n+1)^2} \sum_{-2n \leq j,k \leq 2n} (2n+1-|k|)(2n+1-|j|) |C_{k,j}|^2$$

and by Beppo-Levi's Theorem:

$$\sigma_n^2 \rightarrow \frac{1}{2\pi^2} \|f\|^2$$

The bound $\sup_x |D_n(x)| = O(\log n)$, for the Dirichlet-Kernel ([10], p. 151) gives:

$$\frac{S_m(n)}{n} \leq (\text{const}) \left(\frac{\log n}{n}\right)^{2m} = O(1) \quad (m \geq 3)$$

Moreover $v_{k,j} = 1$ and if we put $v = (2n + 1)^2$

$$\sum_{-n \leq k, j \leq n} v_{k,j}^2 = o(v^{-2}N) \text{ is verified if } n = N^\alpha \text{ with } \alpha < 1/6.$$

Finally if f is periodic continuously differentiable in $C((-\pi, \pi)^2)$ until the second order and it has three derivatives in L^2 with respect to each variable i. e. $\|f_{xxx}\|_2 < \infty$ and $\|f_{yyy}\|_2 < \infty$, then is easy to verify that

$$g(n) = O(n^{-6}) = (v^{1/2}N^{-1}) \text{ if } n = N^\alpha \text{ with } \alpha > \frac{1}{7}.$$

Then, under the above conditions for f , we get

$$W - \lim_{N \rightarrow \infty} \left[\frac{N}{2n + 1} \|\hat{f}_{n,N} - f\|^2 - \frac{2n + 1}{4\pi^2} \right] = N(0, \sigma^2)$$

$$\text{if } n = N^\alpha, \frac{1}{7} < \alpha < 1/6, \text{ and } \sigma^2 = \frac{1}{2\pi^2} \|f\|^2.$$

Note. — This result was obtained by Naradaya [8] in the univariate case, although our method improves the choice of the exponent α . With minor changes the same proof applies to densities on \mathbb{R}^p .

b) As a second example we consider an asymptotic test of goodness of fit for uniform distribution on a sphere.

The basis for $L^2(S^2)$ with the measure invariant by rotations is denoted by $\{Y_n^m(\theta, \phi)\}_{m=-n}^n, n = 0, 1, \dots \left(0 \leq \theta < 2\pi, -\frac{\pi}{2} \leq \phi < \pi/2\right)$ and constructed from the spherical harmonics ([3], p. 511).

We put, with the obvious notations

$$\hat{f}_{n,N}(\theta, \phi) = \sum_{k=0}^n \sum_{m=-k}^k \hat{C}_{m,k} Y_k^m(\theta, \phi)$$

$$\hat{C}_{m,k} = \frac{1}{N} \sum_{i=1}^N Y_k^m(\theta_i, \phi_i)$$

where $(\theta_1, \phi_1), \dots, (\theta_N, \phi_N)$ is the observed sample.

The statistic $T_{n,N} = \|\hat{f}_{n,N} - 1\|_{L^2(S^2)}^2$ can be used to test uniformity.

One easily verifies that $A_n = Id$, so that $\Delta_n = 1$, $\sigma_n^2 = 2$, $\frac{S_m(n)}{n} = 1$.
Moreover

$$\sum_{j=1}^n v_j^2 = O(n^3) = o(v^{-2}N) \quad (\text{with } v = n^2 + 1) \quad \text{for } n = N^\alpha \quad \text{and } \alpha < 1/7.$$

Hence, Theorem 3.2 gives

$$W - \lim_{N \rightarrow \infty} \left[\frac{N}{\sqrt{n^2 + 1}} T_{n,N} - (n^2 + 1)^{1/2} \right] = N(0, 2)$$

c) As a final example, let $f \in L^2(-1, 1)$, $r(x) = 1$ and $\phi_j = (2j + 1)^{1/2} p_j$, p_j the sequence of Legendre polynomials

$$v_j \leq \sqrt{2j + 1}$$

So that

$$\sum_{j=1}^n v_j^2 = O(n^{-2}) = o(n^{-2}N) \quad \text{for } n = N^\alpha \quad \text{and } \alpha < 1/4.$$

The remaining conditions can be verified using the following Theorem ([5], p. 116).

THEOREM 4.1. — With the above notations and $f \in C[-1, 1]$, consider Toeplitz matrices.

$$A_n(f) = \left(\int_{-1}^1 \phi_i(x) \phi_j(x) f(x) dx, \quad i, j = 1, \dots, n \right)$$

If $\lambda_i^{(n)}$ ($i = 1, \dots, n$) are the eigenvalues of $A_n(f)$, then for each $m \geq 1$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (\lambda_i^{(n)})^m = \frac{1}{\pi} \int_{-1}^1 f^m(x) \frac{1}{\sqrt{1-x^2}} dx$$

holds true.

In our case, we get

$$\lim_{n \rightarrow \infty} \sigma_n^2 = \frac{2}{\pi} \int_{-1}^1 f^2(x) \frac{1}{\sqrt{1-x^2}} dx$$

and

$$S_m(n) = \sum_{i=1}^m (\lambda_i^{(n)})^m = O(n).$$

If assume that $f^4 \in C[-1, 1]$, then $g(n) = O(n^{-5})$ (see [6], p. 209) and $g(n) = o(n^{1/2}N^{-1})$ if $n = N^\alpha$, $\alpha > 2/11$.

Summing up, if $f^4 \in C[-1, 1]$, $n = N^\alpha$, $\frac{2}{11} < \alpha < 1/4$, then

$$W - \lim_{N \rightarrow \infty} \left[\frac{N}{n^{1/2}} \| \hat{f}_{n,N} - f \|^2 - n^{1/2} \Delta_n \right] = N(0, \sigma^2)$$

with

$$\sigma^2 = \frac{2}{\pi} \int_{-1}^1 \frac{f^2(x)}{\sqrt{1-x^2}} dx.$$

5. ASYMPTOTIC BEHAVIOUR UNDER CONTIGUOUS ALTERNATIVES

Suppose that one wants to test the null hypothesis

$$H_0 : f = f^0$$

against the sequence of alternatives

$$H_N : f^N(x) = f^0(x) + \delta_N \Phi(x)$$

where Φ is a fixed function in $L^2(\mu)$ and $\delta_N \rightarrow 0$.

The following Theorem states that $T_{n,N} = \frac{N}{n} \| f_{n,M} - f_n \|^2$ is asymptotically gaussian under H_N . The proof follows the lines of [1], th. 4.2 and [8], th. 4.2, and the result can be applied to the previous examples.

We must define before

$$\tilde{\sigma}_n^2 = \frac{2}{n} \sum_{i=1}^n \sum_{j=1}^n \left(\int_{\mathbb{R}^p} \phi_i(x) \phi_j(x) r^2(x) f^N(x) dx \right)^2$$

THEOREM 5.1.— Under H_N , if $\Delta_n = \Delta + o\left(\frac{1}{\sqrt{n}}\right)$, $\tilde{\sigma}_n^2 \rightarrow \sigma_0^2 > 0$, suppose also that the hypothesis of the Theorem 3.2 are satisfied for $n = N^\alpha$, $0 < \alpha < \alpha_0$ and $\delta_N = (N^{-\frac{(2-\alpha)}{4}})$ then

$$W - \lim_{N \rightarrow \infty} \sqrt{n} \left(\frac{T_{n,N} - \Delta}{\sigma_0^2} \right) = N \left(\frac{1}{\sigma_0} \| \Phi \|^2, 1 \right)$$

Proof. — Denote

$$E_{H_N}(\hat{f}_{n,N}) = f_n^N.$$

Where E_{H_N} denotes the expectation when the true underlying distribution has density f^N .

Let $\{\phi_i\}$ a complete orthonormal basis for $L^2(\mu)$ and $\gamma_i = (\Phi, \phi_i)$ the Fourier coefficients of the function Φ . Define

$$\tilde{T}_{n,N} = \frac{N}{n} \|\hat{f}_{n,N} - f_n^N\|^2.$$

We have

$$\|\hat{f}_{n,N} - f_n^0\|^2 = \|\hat{f}_{n,N} - f_n^N\|^2 + 2 \langle \hat{f}_{n,N} - f_n^N, f_n^N - f_n^0 \rangle + \|f_n^N - f_n^0\|^2$$

So that:

$$\begin{aligned} & \sqrt{n} \left(\frac{T_{n,N} - \Delta_0}{\sigma_0} \right) \\ &= \sqrt{n} \left(\frac{\tilde{T}_{n,N} - \Delta_0}{\sigma_0} \right) + \frac{2}{\sqrt{n} \sigma_0} \langle \hat{f}_{n,N} - f_n^N, f_n^N - f_n^0 \rangle + \frac{N}{\sigma_0 \sqrt{n}} \|f_n^N - f_n^0\|^2 \end{aligned}$$

but

$$\frac{N}{\sigma_0 \sqrt{n}} \|f_n^N - f_n^0\|^2 = \frac{\delta_N^2 N}{\sigma_0 \sqrt{n}} \|\Phi_n\|^2 \xrightarrow{n \rightarrow \infty} \frac{\|\Phi\|^2}{\sigma_0}$$

moreover

$$E \left[\frac{2}{\sqrt{n} \sigma_0} \langle \hat{f}_{n,N} - f_n^N, f_n^N - f_n^0 \rangle \right] = 0.$$

and

$$\begin{aligned} E \left[\frac{2}{\sqrt{n} \sigma_0} \langle \hat{f}_{n,N} - f_n^N, f_n^N - f_n^0 \rangle \right]^2 &= \frac{4N\delta_N^2}{n\sigma_0^2} E(\sqrt{N}(\hat{f}_{n,N} - f_n^N), \Phi_n)^2 \\ &= \frac{4N\delta_N^2}{\sqrt{n}\sigma_0^2} \frac{1}{\sqrt{n}} \left(\sum_{i=1}^n \sum_{j=1}^n \left(\int \phi_i \phi_j r^2 f^N dx - a_i^N a_j^N \right) \gamma_i \gamma_j \right) \\ &\leq k \frac{1}{\sqrt{n}} \|rf_N\|_\infty \|\Phi_n\|^2 \leq k \frac{1}{\sqrt{n}} \|rf\|_\infty \|\Phi\|^2 \end{aligned}$$

then we can conclude

$$\frac{2}{\sqrt{n} \sigma_0} \langle \hat{f}_{n,N} - f_n^N, f_n^N - f_n^0 \rangle \xrightarrow{N \rightarrow \infty} 0$$

and finally applying Theorem 3.2

$$W - \lim_{N \rightarrow \infty} \left(\frac{\tilde{T}_{n,N} - \Delta}{\sigma_0} \right) = N(0, 1).$$

ACKNOWLEDGMENT

The author wishes to express his gratitude to Professor J. Bretagnolle for a careful reading of a first version of this paper, and for his valuable suggestions that improved considerably the results.

REFERENCES

- [1] P. J. BICKEL and M. ROSENBLATT, On some measures of the deviations of density functions estimates. *The Ann. of Statistics*, t. 1, n° 6, 1973, p. 1071-1095.
- [2] N. N. CENCOV, Evaluations of an unknown distribution density from observations. *Soviet Math.*, t. 3, 1962, p. 1559-1562.
- [3] R. COURANT and D. HILBERT, *Methods of Mathematical Physics*, t. 1, New York, 1953.
- [4] E. GINÉ, J. LEÓN, On the central limit theorem in Hilbert Spaces. *Stochastica*, t. IV, n° 1, 1980, p. 43-71.
- [5] U. GRENANDER, Szegő. Toeplitz forms and their application. University of California Press. Berkeley. 1958.
- [6] E. ISACCON, H. KELLER, « Analysis of Numerical Methods » John Wiley and Sons. Inc. New York, 1966.
- [7] KUELBS, KURTZ, Bery Essen estimates in Hilbert Space and an application to the LIL. *Ann. Probability*, t. 2, 1974, p. 387-407.
- [8] E. A. NARADAYA, On a quadratic measure of deviation of the estimate of a distribution density. *Theo. Prob. and its App.*, n° 4, 1976, p. 844-850.
- [9] A. ZYGMUND, Trygonometric series, 2nd Ed. Cambridge University Press. New York, 1959.

(Manuscrit reçu le 15 octobre 1982)