NONPARAMETRIC DENSITY ESTIMATION OF CONTINUOUS PART OF A MIXED MEASURE

Rachid Sabre

1. INTRODUCTION

This paper deals with the estimation of the bivariate density of the continuous part of a certain mixture. More precisely, we consider a pair of the random variables (X,Y) whose probability measure is the sum of an absolutely continuous measure with respect to the Lebesgue measure, a discrete measure and a finite number of absolutely continuous measures on several lines:

$$d\mu = f(x,y)dx\,dy + \sum_{j=1}^{q} a'_j \delta_{(w_{1j},w_{2j})} + \sum_{i=1}^{q'} \varphi_i(u_1)\delta_{(u_1,a_iu_1+b_i)}, \quad u_1 \in \mathbb{R} \tag{1}$$

where the numbers q and q' are supposed nonnegative integers and known. f is the density of the continuous variable which is supposed to be a nonnegative uniformly continuous function. The real positive number a'_j is the amplitude of the jump at (w_{1j},w_{2j}) and is assumed unknown. The densities φ_i are nonnegative uniformly continuous functions assumed unknown. The coefficients of the lines a_i, b_i are real numbers assumed known. δ is the Dirac measure. The jump points (w_{1j},w_{2j}) are known real numbers. The theorem 3.1 gives an estimator which can be used to verify the existence of jump at any point, see the remark 3.1. However, in an experimental way, we suggest an intuitive technique for localize the jump point (w_{1j},w_{2j}) in a block $[\alpha_{1j},\beta_{1j}]\times[\alpha_{2j},\beta_{2j}]$. Indeed, we calculate the empirical distribution for several samples of (X,Y) and if we remark for different samples the presence of a jump at points close each other, we give therefore a block containing this jump point, obviously this block depends on the number of samples taken. The block is assumed sufficiently small to contain only one jump point.

A concrete example concerns the study of structural fissure of the agricultural soil. On a homogenous soil, measures of the resistance variable X and the humidity variable Y are taken at several locations at a depth of a 30cm. The measurement values are distributed according to a Gaussian law, except in certain locations where the experimentalist find small galleries where measurement values of resistance and humidity decrease (the presence of jumps). When the measures are made in places where the passage of tractors is frequent, the variable Y becomes linear with respect to the variable X and their measurers follow a new distribution noted φ_i (the presence of some measures continuous on the lines determined by the frequent passages of tractors). In this case we will consider the model (1).

The goal of this work is to estimate, for every real pair (x, y), the density $f(x, y)$, from a sample with a finite size of the random variables (X,Y). Indeed, when (x, y) satisfies $x \notin [\alpha_{1j}, \beta_{1j}]$, $y \notin [\alpha_{2j}, \beta_{2j}]$ and $y \neq a_i x + b_i$, we use the classical kernel estimate as in (Parzen, 1962), (Rosemblatt, 1965), (Bosq and Lecoutre, 1987) and (Deheuvels, 1977, 1979, 1980). For the other points, in order to obtain an asymptotically unbiased and consistent estimate, we smooth the kernel estimate by using four windows satisfying some conditions. The same technique is used in (Sabre, 1994, 1995) to estimate the spectral density function. We give an estimator $\hat{a}'_j(x, y)$ converging to the amplitude a'_j if $(x, y) = (w_{1j}, w_{2j})$ or to zero otherwise. Thus, we have an estimate of the amplitude of jump point when this jump point is exactly known. We can use this result to verify the presence of the jump at any point (x, y). We give an asymptotically unbiased and consistent estimate of the density φ_i.

Theoretically, our work is true for all q and q' real numbers. Because it is not always easy to determine the blocks containing the jump points and the fact that we propose different estimates with respect to the location of the point, our estimation is interesting where q and q' are small.

We conclude this paper by considering and studying the simulation of the particular case where we have one random variable X whose probability measure μ is a sum of an absolutely continuous measure with respect to the Lebesgue measure and a discrete measure:

$$d\mu = f(x)dx + \sum_{m=1}^{q} a_m \delta_{\lambda_m} \tag{2}$$

As in the two-dimensional case, we smooth the kernel estimate by using two windows satisfying certain conditions. Thus we give an asymptotically unbiased and consistent estimate of the density function f.

The motivation of this work is that, in practice, it often occurs that the observed data have the same distribution as the one of a usual law except in some points where we have a discontinuity of the law observed. In this case we can consider that the law of data observed is the sum of the usual law with another discrete law. It is therefore interesting to estimate the density of the continuous part, especially at jump points.

For example when we consider the regression model, $Y = h(X) + \varepsilon$, ε must be a centered Gaussian variable. To show that we take a sample of the residues $\varepsilon_i = Y_i - h(X_i)$ $i = 1, 2, ...n$. If the empirical distribution of residues is Gaussian except at discrete points, for example λ_1 and λ_2, the law of ε is: $d\mu = f(x)dx + a_1\delta_{\lambda_1} + a_2\delta_{\lambda_2}$. If we show that f is the density of a Gaussian centered variable, we are sure that we must change our model by adding an other discrete variable.

A concrete application of the one-dimensional case deals with the process of filling bottles of a 33cl volume each. To control the quality of this process, we check that this process is evenly distributed according to a Gaussian distribution. The experimentalist has taken a sample of 115 bottles and measured the quantity of liquid contained in these bottles. The measures usually gives about 33cl. Due to an abnormal disfunction caused by a random slowing or acceleration of the motion of the rolling band. Thus the measure randomly increases or decrease to reach a constant.

We briefly indicate the organization of this paper: In the second section we consider a pair of random variables defined by (1). First, we estimate the density f outside the neighbourhood of the jump point (theorem 2.1) and we study the estimation of the density f inside the neighbourhood of a jump point (theorems 2.2, 2.3). In the third section we estimate the amplitude of the jump points and the densities on several lines (theorem 3.1 and theorem 3.2). The fourth section provides the proofs the theorems. We finish by studying the simulation of the estimate for the univariate case.

2. KERNEL ESTIMATE OF THE DENSITY FUNCTION

In this section we consider a pair of random variables (X,Y) whose probability measure, μ, is defined by (1). Our goal is to estimate, for every real pair (x, y), the density function f.

Notation:

$$\boldsymbol{B} = \{(x, y) \in \mathbb{R}^2 \quad \text{such that} \quad \exists i \in \{1, ..., q'\} : y = a_i x + b_i\}$$

$$\boldsymbol{A} = \bigcup_{j=1}^{q} ([\alpha_{1j}, \beta_{1j}] \times \mathbb{R}) \cup (\mathbb{R} \times [\alpha_{2j}, \beta_{2j}]) \cup \boldsymbol{B}.$$

Let $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ independent observations of (X,Y). To estimate the density function f at point (x, y) we distinguish two cases.

2.1. *The estimation of the function f outside $\boldsymbol{A}$*

Let (x, y) be a pair outside $\boldsymbol{A}$, we consider the kernel estimate defined as follows:

$$\widehat{f_n}(x,y)=\sum_{i=1}^{n}\frac{1}{nh_n^2}K\left(\frac{x-x_i}{h_n},\frac{y-y_i}{h_n}\right) \tag{3}$$

where K is defined by $K(u,v)=K_1(u)K_2(v)$ with K_1 and K_2 two continuous, even, decreasing kernels such that: $\int\left\|y^2\right\|K_i(y)dy<\infty \quad i=1,2$. The smoothing parameter h_n, converges to zero and nh_n^2 converges to the infinite.

First we show that $\widehat{f_n}$ is an asymptotically unbiased and consistent estimate of f outside $\boldsymbol{A}$, we assume that $\frac{1}{h_n}K_1\left(\frac{1}{h_n}\right)$ and $\frac{1}{h_n}K_2\left(\frac{1}{h_n}\right)$ converge to zero, for example $K_1(x)=\frac{1}{(2\pi)^{\frac{1}{2}}}\exp\left(\frac{-x^2}{2}\right)$.

Theorem 2.1. Let (x,y) be a pair outside $\boldsymbol{A}$, then $\widehat{f_n}(x,y)$ is an asymptotically unbiased and consistent estimate. If f is twice differentiable and its partial derivatives are continuous and bounded, then

$$\mathrm{E}(\widehat{f_n}(x,y))-f(x,y)=O(h_n^2)+O\left(\frac{1}{h_n}K_1\left(\frac{1}{h_n}\right)\right)+O\left(\frac{1}{h_n}K_2\left(\frac{1}{h_n}\right)\right) \tag{4}$$

$$\mathrm{Var}(\widehat{f_n}(x,y))=O\left(\frac{1}{nh_n^2}\right)+O\left(\frac{1}{nh_n^3}K_1^2\left(\frac{1}{h_n}\right)\right)+O\left(\frac{1}{nh_n^3}K_2^2\left(\frac{1}{h_n}\right)\right) \tag{5}$$

2.2. *Estimation of the function* f *inside* $\boldsymbol{A}$

In order to estimate the density $f(x,y)$ where $(x,y)\in\boldsymbol{A}$, we smooth the kernel estimate $\widehat{f_n}$ by using four windows defined as follows:

$$W_n^{(1)}(t)=M_n^{(1)}W^{(1)}(tM_n^{(1)})\,;\; W_n^{(2)}(t)=M_n^{(2)}W^{(2)}(tM_n^{(2)})\,;$$

$$W_n^{(3)}(t)=L_n^{(1)}W^{(3)}(tL_n^{(1)}) \text{ et } W_n^{(4)}(t)=L_n^{(2)}W^{(4)}(tL_n^{(2)})$$

where $M_n^{(1)}$, $M_n^{(2)}$, $L_n^{(1)}$ and $L_n^{(2)}$ are nonnegative real sequences satisfying:

$$M_n^{(r)}\to+\infty\,;\; L_n^{(r)}\to+\infty\,;\; M_n^{(r)}h_n\to 0\,;\; L_n^{(r)}h_n\to 0\,;$$

where h_n is defined in (3). $W^{(i)}$ is a nonnegative, even, integrable function vanishing outside the interval $[-1,1]$ such that $\int_{-1}^{1} W^{(i)}(x)dx = 1$, $i = 1,2,3,4$ and moreover satisfying the following equalities:

$$W^{(2)}(M_n^{(2)}\theta) - W^{(1)}(M_n^{(1)}\theta) = 0 \qquad \forall \theta \in \left]\frac{-1}{M_n^{(1)}}, \frac{1}{M_n^{(1)}}\right[. \tag{6}$$

$$W_n^{(4)}(L_n^{(2)}\theta) - W^{(3)}(L_n^{(1)}\theta) = 0 \qquad \forall \theta \in \left]\frac{-1}{L_n^{(1)}}, \frac{1}{L_n^{(1)}}\right[. \tag{7}$$

Now we consider the estimate $\widehat{g_n}$ defined by:

$\widehat{g_n}(x,y) = \int_{\mathbb{R}^2} \mathrm{S}_n(x-u_1)\mathrm{R}_n(y-u_2)\widehat{f_n}(u_1,u_2)du_1du_2$, where

$$\mathrm{S}_n(v) = \frac{W_n^{(2)}(v) - \dfrac{M_n^{(2)}}{M_n^{(1)}} W_n^{(1)}(v)}{1 - \dfrac{M_n^{(2)}}{M_n^{(1)}}}$$

$$\mathrm{R}_n(v) = \frac{W_n^{(3)}(v) - \dfrac{L_n^{(2)}}{L_n^{(1)}} W_n^{(4)}(v)}{1 - \dfrac{L_n^{(2)}}{L_n^{(1)}}}$$

and $\widehat{f_n}$ is defined in (3) we assume that $\frac{1}{h_n}K_1\left(\frac{1}{h_n M_n^{(1)}}\right)$ and $\frac{1}{h_n}K_2\left(\frac{1}{h_n L_n^{(1)}}\right)$ converge to zero, for example $K_1(x) = \frac{1}{(2\pi)^{\frac{1}{2}}}\exp\left(\frac{-x^2}{2}\right)$.

We show in the following theorem that $\widehat{g_n}(x,y)$ is an asymptotically unbiased estimate of f on the set $\boldsymbol{A}$ and we give the rate of convergence of the bias.

Theorem 2.2. Let (x,y) be an element of $\boldsymbol{A}$. Then $\widehat{g_n}(x,y)$ is an asymptotically unbiased estimate. If f is twice differentiable and its partial derivatives are bounded, we have:

$$\mathrm{E}(\widehat{g_n}(x,y))-f(x,y)=O(h_n^2)+O\left(\frac{1}{h_n}K_1\left(\frac{1}{h_n}\right)+\frac{1}{h_n}K_2\left(\frac{1}{h_n}\right)\right)+$$

$$\begin{cases} O\left(\frac{1}{h_n^2}K_1\left(\frac{1}{M_n^{(1)}h_n}\right)K_2\left(\frac{1}{L_n^{(1)}h_n}\right)\right) & \text{if } \begin{cases}\exists \mathrm{j}\in\{1,\dots,\mathrm{q}\}: \\ (x,y)\in[\alpha_{1j},\beta_{1j}]\times[\alpha_{2j},\beta_{2j}]\end{cases} \\ O\left(\frac{1}{h_n^2}K_1\left(\frac{1}{M_n^{(1)}h_n}\right)K_2\left(\frac{1}{h_n}\right)\right) & \text{if } \begin{cases}\exists \mathrm{j}\in\{1,\dots,\mathrm{q}\}: \\ x\in[\alpha_{1j},\beta_{1j}] \text{ and } y\notin[\alpha_{2j},\beta_{2j}]\end{cases} \\ O\left(\frac{1}{h_n^2}K_1\left(\frac{1}{h_n}\right)K_2\left(\frac{1}{L_n^{(1)}h_n}\right)\right) & \text{if } \begin{cases}\exists \mathrm{j}\in\{1,\dots,\mathrm{q}\}: \\ x\notin[\alpha_{1j},\beta_{1j}] \text{ and } y\in[\alpha_{2j},\beta_{2j}]\end{cases} \\ O\left(\frac{1}{h_n}K_1\left(\frac{1}{h_n}\right)+\frac{1}{h_n}K_2\left(\frac{1}{h_n}\right)\right) & \text{if } \begin{cases}\forall j\in\{1,\dots,q\}: \\ x\notin[\alpha_{1j},\beta_{1j}];\ y\notin[\alpha_{2j},\beta_{2j}] \text{ and} \\ \exists i\in\{1,\dots,q\}: y=a_i x+b_i\end{cases} \end{cases}$$

We show in the following theorem that $\widehat{g_n}(x,y)$ is a consistent estimate of the function f on the set $\boldsymbol{A}$, thus giving the rate of convergence of its variance.

Theorem 2.3. Let (x,y) belong to $\boldsymbol{A}$. Then $\widehat{g_n}(x,y)$ is a consistent estimator. If f is twice differentiable and its partial derivatives are bounded, then

$$\mathrm{var}(\widehat{g_n}(x,y))=O\left(\frac{1}{nh_n^2}\right)+O\left(\frac{1}{nh_n^3}K_1^2\left(\frac{1}{h_n}\right)\right)+O\left(\frac{1}{nh_n^3}K_2^2\left(\frac{1}{h_n}\right)\right)+$$

$$\begin{cases} O\left(\frac{1}{nh_n^4}K_1\left(\frac{1}{M_n^{(1)}h_n}\right)K_2\left(\frac{1}{L_n^{(1)}h_n}\right)\right) & \text{if } \begin{cases}\exists \mathrm{j}\in\{1,\dots,\mathrm{q}\}: \\ (x,y)=(w_{1j},w_{2j})\end{cases} \\ O\left(\frac{1}{nh_n^4}K_1\left(\frac{1}{M_n^{(1)}h_n}\right)K_2\left(\frac{1}{h_n}\right)\right) & \text{if } \begin{cases}\exists \mathrm{j}\in\{1,\dots,\mathrm{q}\}: \\ x=w_{1j} \text{ and } y\neq w_{2j}\end{cases} \\ O\left(\frac{1}{nh_n^4}K_1\left(\frac{1}{h_n}\right)K_2\left(\frac{1}{L_n^{(1)}h_n}\right)\right) & \text{if } \begin{cases}\exists \mathrm{j}\in\{1,\dots,\mathrm{q}\}: \\ x\neq w_{1j} \text{ and } y=w_{2j}\end{cases} \\ 0 & \text{if } \begin{cases}\forall j\in\{1,\dots,q\}: \\ x\neq w_{1j};\ y\neq w_{2j} \\ \text{and } \exists i\in\{1,\dots,q'\}: \\ y=a_i x+b_i\end{cases} \end{cases}$$

2.3. *Univariate case*

In the univariate case, we consider one random variable X whose probability measure, μ, is the sum of an absolutely continuous measure with respect to the Lebesgue measure and a discrete measure, defined in (2).

The estimate of density function f can be given as a particular case of the two-dimensional case, namely

$$\hat{h}(x)=\begin{cases}\widehat{g_n}(x)=\int_{\mathbb{R}} S_n(x-y)\widehat{f_n}(y)dy & \text{if } x\in \boldsymbol{A}\\ \widehat{f_n}(x) & \text{if } x\notin \boldsymbol{A}\end{cases}$$

where $\widehat{f_n}(x)=\sum_{i=1}^{n}\frac{1}{nh_n}K\left(\frac{x-x_i}{h_n}\right)$ and $S_n(z)=\dfrac{W_n^{(2)}(z)-\frac{M_n^{(2)}}{M_n^{(1)}}W_n^{(1)}(z)}{1-\frac{M_n^{(2)}}{M_n^{(1)}}}$ and

$\boldsymbol{A}=\bigcup_{m=1}^{q}[\alpha_m,\beta_m]$ the intervals contain the jump points $(\lambda_m\in[\alpha_m,\beta_m])$. The windows are defined as follows:

$$W_n^{(1)}(t)=M_n^{(1)}W^{(1)}(tM_n^{(1)})\,;W_n^{(2)}(t)=M_n^{(2)}W^{(2)}(tM_n^{(2)});$$

where $M_n^{(1)}$ and $M_n^{(2)}$ are nonnegative real sequences satisfying:

$$M_n^{(r)}\to+\infty\,;\; M_n^{(r)}h_n\to 0$$

where h_n is defined in (3). $W^{(i)}$ is a nonnegative, even, integrable function vanishing outside the interval $[-1,1]$ such that $\int_{-1}^{1}W^{(i)}(x)dx=1$, i = 1, 2 and moreover satisfying the following equalities:

$$W^{(2)}(M_n^{(2)}\theta)-W^{(1)}(M_n^{(1)}\theta)=0 \qquad \forall\theta\in\left]\frac{-1}{M_n^{(1)}},\frac{1}{M_n^{(1)}}\right[.$$

3. THE ESTIMATION OF THE AMPLITUDE a'_j OF THE JUMP AND THE DENSITIES φ_i

In this section we purpose the following estimator:

$$\widehat{a'}_n(x,y)=\frac{h_n^2}{K(0,0)}\sum_{i=1}^{n}\frac{1}{nh_n^2}K\left(\frac{x-x_i}{h_n},\frac{y-y_i}{h_n}\right)$$

where K is defined in (3) such that $K(0,0)\neq 0$.

Theorem 3.1. We have

$$\mathrm{E}(\widehat{a'}_n(x,y)) = \begin{cases} O(h_n) & \text{if } (x,y) \neq (w_{1j}, w_{2j}) \quad \forall \mathrm{j} \in \{1,2,...,\mathrm{q}\} \\ a'_j + O(h_n^2) & \text{if } (x,y) = (w_{1j}, w_{2j}) \quad \mathrm{j} \in \{1,2,...,\mathrm{q}\} \end{cases}$$

$$\mathrm{Var}(\widehat{a'}_n(x,y)) = O\left(\frac{1}{n}\right).$$

Remark 3.1. This result can be used to prove the presence of jump at point (x,y). Indeed, we calculate the empirical mean of $\widehat{a'}_n(x,y)$ from several samples. If its value is approximatively zero we consider that there not a jump at the point (x,y). If not there exists a jump at (x,y).

In the following theorem we give an asymptotically unbiased and consistent estimate of the density φ_i.

Let λ_1 a real number and λ_2 is such that $\lambda_2 = a_i\lambda_1 + b_i$ where $i \in \{1,2,...,q'\}$. We estimate $\varphi_i(\lambda_1)$ by

$$\hat{\varphi}_i(\lambda_1) = \frac{1}{K_2(0)} \sum_{i=1}^{n} \frac{1}{nh_n h'_n} K\left(\frac{\lambda_1 - x_i}{h_n}, \frac{\lambda_2 - y_i}{h'_n}\right)$$

where $h'_n \to 0$ and $\frac{h_n}{h'_n} \to 0$ and $nh'_n \to \infty$ and $nh_n^2 h'_n \to \infty$.

Theorem 3.2. We have

$$\mathrm{E}\hat{\varphi}_i(\lambda_1) - \varphi_i(\lambda_1) = O(h_n) + O\left(\frac{1}{h_n} K\left(\frac{1}{h_n}, \frac{1}{h'_n}\right)\right)$$

$$\mathrm{Var}(\hat{\varphi}_i(\lambda_1)) = O\left(\frac{1}{nh_n^2 h'_n}\right).$$

4. PROOFS

4.1. *Proof of the theorem 2.1*

From (1) we have

$$\mathrm{E}(\widehat{f_n}(x,y)) = \frac{1}{h_n^2}\int K\left(\frac{x-z_1}{h_n},\frac{y-z_2}{h_n}\right) f(z_1,z_2)dz_1dz_2$$
$$+\frac{1}{h_n^2}\sum_{j=1}^{q} a'_j\, K\left(\frac{x-w_{1j}}{h_n},\frac{y-w_{2j}}{h_n}\right)$$
$$+\sum_{j=1}^{q'}\frac{1}{h_n^2}\int K\left(\frac{x-v_1}{h_n},\frac{y-a_iv_1-b_i}{h_n}\right)\varphi_i(v_1)dv_1$$

We note respectively the 3 terms of the last equality by A_n, B_n and C_n. We know, from the works of (Parzen, 1962), (Rosemblatt, 1965) and (Bosq and Lecoutre, 1987), that $A_n - f(x,y) = O(h_n^2)$. On the other hand as $(x,y) \neq (w_{j1}, w_{j2})$, we have

$$B_n = O\left(\frac{1}{h_n^2}K_1\left(\frac{1}{h_n}\right)K_2\left(\frac{1}{h_n}\right)\right). \tag{8}$$

Let us now show that C_n tends to 0. Indeed, we assume that $x < \frac{y-b_i}{a_i}$ (same arguments in the case where $x > \frac{y-b_i}{a_i}$) and we split the integral, in the expression of C_n as follows:

$$C_n = \frac{1}{h_n^2}\sum_{i=1}^{q'}\int_{-\infty}^{x-\varepsilon} K_1\left(\frac{x-v_1}{h_n}\right)K_2\left(\frac{y-a_iv_1-b_i}{h_n}\right)\varphi_i(v_1)dv_1$$
$$+\frac{1}{h_n^2}\sum_{i=1}^{q'}\int_{x-\varepsilon}^{x+\varepsilon} K_1\left(\frac{x-v_1}{h_n}\right)K_2\left(\frac{y-a_iv_1-b_i}{h_n}\right)\varphi_i(v_1)dv_1$$
$$+\frac{1}{h_n^2}\sum_{i=1}^{q'}\int_{x+\varepsilon}^{\frac{y-b_i}{a_i}-\varepsilon} K_1\left(\frac{x-v_1}{h_n}\right)K_2\left(\frac{y-a_iv_1-b_i}{h_n}\right)\varphi_i(v_1)dv_1$$
$$+\frac{1}{h_n^2}\sum_{i=1}^{q'}\int_{\frac{y-b_i}{a_i}-\varepsilon}^{\frac{y-b_i}{a_i}+\varepsilon} K_1\left(\frac{x-v_1}{h_n}\right)K_2\left(\frac{y-a_iv_1-b_i}{h_n}\right)\varphi_i(v_1)dv_1$$
$$+\frac{1}{h_n^2}\sum_{i=1}^{q'}\int_{\frac{y-b_i}{a_i}+\varepsilon}^{+\infty} K_1\left(\frac{x-v_1}{h_n}\right)K_2\left(\frac{y-a_iv_1-b_i}{h_n}\right)\varphi_i(v_1)dv_1,$$

where ε is a nonnegative real sufficiently small for having $x+\varepsilon < \frac{y-b_i}{a_i} - \varepsilon$. We note the five terms of the last equality by I_1, I_2, I_3, I_4 and I_5. Since the functions K_1 and K_2 are decreasing and even, we can write

$$I_1 \le \frac{1}{h_n^2} \sup_{v_1 \in]-\infty, x-\varepsilon[} K_1\left(\frac{x-v_1}{h_n}\right) \sup_{v_1 \in]-\infty, x-\varepsilon[} K_2\left(\frac{y-a_i v_1 - b_i}{h_n}\right) \int_{-\infty}^{+\infty} \varphi_i(v_1) dv_1 .$$

The two 'sup' reach values respectively different from x and from $\frac{y-b_i}{a_i}$, hence

$$I_1 = O\left(\frac{1}{h_n^2} K_1\left(\frac{1}{h_n}\right) K_2\left(\frac{1}{h_n}\right)\right). \tag{9}$$

as in above, it is shown that

$$I_3 = O\left(\frac{1}{h_n^2} K_1\left(\frac{1}{h_n}\right) K_2\left(\frac{1}{h_n}\right)\right) \text{ and } I_5 = O\left(\frac{1}{h_n^2} K_1\left(\frac{1}{h_n}\right) K_2\left(\frac{1}{h_n}\right)\right).$$

On the other hand for all v belonging to $[x-\varepsilon, x+\varepsilon]$ we have $y \neq a_i v - b_i$. Therefore we have

$$I_2 \le \frac{1}{h_n^2} \sup_{v \in [x-\varepsilon, x+\varepsilon]} K_2\left(\frac{y-a_i v - b_i}{h_n}\right) \int_{-\infty}^{+\infty} K_1\left(\frac{x-v_1}{h_n}\right) \varphi_i(v_1) dv_1 .$$

Since $x \to \frac{1}{h_n} K_1\left(\frac{x}{h_n}\right)$ is a kernel, we conclude that $I_2 = O\left(\frac{1}{h_n} K_2\left(\frac{1}{h_n}\right)\right)$. In the same manner we increase the expression of I_4. Thus we obtain

$$C_n = O\left(\frac{1}{h_n} K_1\left(\frac{1}{h_n}\right)\right) + O\left(\frac{1}{h_n} K_2\left(\frac{1}{h_n}\right)\right). \tag{10}$$

From (1) and (3) we have $\operatorname{Var} \widehat{f_n}(x, y) = H_1 + H_2 + H_3 - \frac{1}{n} \mathrm{E}^2(\widehat{f_n}(x, y))$, where

$$H_1 = \frac{1}{nh_n^4} \int K^2\left(\frac{x-z_1}{h_n}, \frac{y-z_2}{h_n}\right) f(z_1, z_2) dz_1 dz_2 ,$$

$$H_2 = \frac{1}{nh_n^4} \sum_{j=1}^{q} a_j K^2\left(\frac{x-w_{1j}}{h_n}, \frac{y-w_{2j}}{h_n}\right),$$

$$H_3 = \frac{1}{nh_n^4} \int \sum_{i=1}^{q'} K^2\left(\frac{x-u}{h_n}, \frac{y-a_i u - b_i}{h_n}\right) \varphi_i(u) du .$$

putting $x - z_1 = h_n t_1$ and $x - z_2 = h_n t_2$ in the integral of H_1, we obtain

$$H_1 = \frac{1}{nh_n^2} \int K^2(t_1, t_2) f(t_1 h_n - x, t_2 h_n - y) dt_1 dt_2$$

From the theorem of the finite increments, we have

$$H_1 \le \frac{1}{nh_n^2} \int K^2(t_1, t_2) \max(f') \|(t_1 h_n, t_2 h_n)\| dt_1 dt_2$$
$$+ \frac{1}{nh_n^2} \int K^2(t_1, t_2) f(x, y) dt_1 dt_2$$

Thus, we obtain $H_1 = O\left(\frac{1}{nh_n}\right) + O\left(\frac{f(x, y)}{nh_n^2} \int\int K^2\right)$. On the other hand as $x \ne w_{1j}$ and $y \ne w_{2j}$, we have $H_2 = O\left(\frac{1}{nh_n^4} K^2\left(\frac{1}{h_n}, \frac{1}{h_n}\right)\right)$.

We write $H_3 = \sum_{i=1}^{q'} P_i$ and we split this integral as follows:

$$P_i = \frac{1}{nh_n^4}\left[\int_{-\infty}^{x-\varepsilon} + \int_{x-\varepsilon}^{x+\varepsilon} + \int_{x+\varepsilon}^{\frac{y-b_i}{a_i}-\varepsilon} + \int_{\frac{y-b_i}{a_i}-\varepsilon}^{\frac{y-b_i}{a_i}+\varepsilon} + \int_{\frac{y-b_i}{a_i}+\varepsilon}^{+\infty} \right]$$

where ε is a nonnegative real number satisfying the following inequality: $x + \varepsilon < \frac{y - b_i}{a_i} - \varepsilon$. We note respectively the five integrals of P_i by P_{i2}, P_{i3}, P_{i4}, and P_{i5}. Similar to the proof of the equality (9) we have:

$P_{i1} = P_{i3} = P_{i5} = O\left(\frac{1}{nh_n^4} K_1^2\left(\frac{1}{h_n}\right) K_2^2\left(\frac{1}{h_n}\right)\right)$. The amount P_{i2} is bounded by $\frac{1}{nh_n^3} \sup_{u \in [x-\varepsilon, x+\varepsilon]} K_2^2\left(\frac{y - a_i u - b_i}{h_n}\right) \sup(K_1) \int_{-\infty}^{+\infty} \frac{1}{h_n} K_1\left(\frac{x-u}{h_n}\right) \varphi_i(u) du$ since K_i is a kernel function, we then have

$$P_{i2} = O\left(\frac{1}{nh_n^3} K_1^2\left(\frac{1}{h_n}\right)\right). \tag{11}$$

In the same manner, we increase the expression of P_{i4} and we obtain

$$P_{i4} = O\left(\frac{1}{nh_n^3}K_1^2\left(\frac{1}{h_n}\right)\right). \tag{12}$$

Consequently

$$P_i = O\left(\frac{1}{nh_n^3}K_1^2\left(\frac{1}{h_n}\right)\right) + O\left(\frac{1}{nh_n^3}K_2^2\left(\frac{1}{h_n}\right)\right). \tag{13}$$

Thus we deduce the equality (5).

4.2. *Proof of the theorem 2.2*

For a large n, we have

$$\mathrm{E}[\widehat{g_n}(x,y)] - f(x,y) = \int_{\mathbb{R}^2} \mathrm{S}_n(x-u_1)\mathrm{R}_n(y-u_2)(\mathrm{E}[\widehat{f_n}(u_1,u_2)] - f(x,y))du_1du_2.$$

Split this integral to 9 integrals as follows:

$$\begin{aligned}\mathrm{E}[\widehat{g_n}(x,y)] - f(x,y) = & \int_{-\infty}^{x-\frac{1}{M_n^{(1)}}}\int_{-\infty}^{y-\frac{1}{L_n^{(1)}}} + \int_{-\infty}^{x-\frac{1}{M_n^{(1)}}}\int_{y-\frac{1}{L_n^{(1)}}}^{y+\frac{1}{L_n^{(1)}}} + \int_{-\infty}^{x-\frac{1}{M_n^{(1)}}}\int_{y+\frac{1}{L_n^{(1)}}}^{+\infty} \\ & + \int_{x-\frac{1}{M_n^{(1)}}}^{x+\frac{1}{M_n^{(1)}}}\int_{-\infty}^{y-\frac{1}{L_n^{(1)}}} + \int_{x-\frac{1}{M_n^{(1)}}}^{x+\frac{1}{M_n^{(1)}}}\int_{y-\frac{1}{L_n^{(1)}}}^{y+\frac{1}{L_n^{(1)}}} + \int_{x-\frac{1}{M_n^{(1)}}}^{x+\frac{1}{M_n^{(1)}}}\int_{y+\frac{1}{L_n^{(1)}}}^{+\infty} \\ & + \int_{x+\frac{1}{M_n^{(1)}}}^{+\infty}\int_{-\infty}^{y-\frac{1}{L_n^{(1)}}} + \int_{x+\frac{1}{M_n^{(1)}}}^{+\infty}\int_{y-\frac{1}{L_n^{(1)}}}^{y+\frac{1}{L_n^{(1)}}} + \int_{x+\frac{1}{M_n^{(1)}}}^{+\infty}\int_{y+\frac{1}{L_n^{(1)}}}^{+\infty}\end{aligned}$$

We note these integrals: E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8 and E_9. From (11) we get that E_2, E_4, E_5, E_6 and E_8 are null. Let us show that E_1, E_3, E_7, and E_9 tend to 0. Indeed, by putting $v_1 = M_n^{(2)}(x-u_1)$ and $v_2 = L_n^{(2)}(y-u_2)$ in the integral E_1, and using the fact that $W^{(i)}$ is vanishing outside the interval $[-1,1]$, we obtain

$$E_1 = \frac{1}{\left(1-\frac{M_n^{(2)}}{M_n^{(1)}}\right)\left(1-\frac{L_n^{(2)}}{L_n^{(1)}}\right)}\int_{\frac{M_n^{(2)}}{M_n^{(1)}}}^{1}\int_{\frac{L_n^{(2)}}{L_n^{(1)}}}^{1} W^{(2)}(v_1)W^{(4)}(v_2)\times$$

$$\left[\mathrm{E}\widehat{f_n}\left(x-\frac{v_1}{M_n^{(2)}}, y-\frac{v_2}{L_n^{(2)}}\right) - f(x,y)\right]dv_1dv_2.$$

In the same way, by using the fact that $W^{(i)}$ is even, we get the followings:

$$E_3 = \frac{1}{\left(1-\frac{M_n^{(2)}}{M_n^{(1)}}\right)\left(1-\frac{L_n^{(2)}}{L_n^{(1)}}\right)} \int_{\frac{M_n^{(2)}}{M_n^{(1)}}}^{+\infty} \int_{\frac{L_n^{(2)}}{L_n^{(1)}}}^{+\infty} W^{(2)}(v_1) W^{(4)}(v_2) \times \left[\mathrm{E}\widehat{f}_n\left(x - \frac{v_1}{M_n^{(2)}}, y + \frac{v_2}{L_n^{(2)}}\right) - f(x,y) \right] dv_1 dv_2.$$

$$E_7 = \frac{1}{\left(1-\frac{M_n^{(2)}}{M_n^{(1)}}\right)\left(1-\frac{L_n^{(2)}}{L_n^{(1)}}\right)} \int_{\frac{M_n^{(2)}}{M_n^{(1)}}}^{1} \int_{\frac{L_n^{(2)}}{L_n^{(1)}}}^{1} W^{(2)}(v_1) W^{(4)}(v_2) \times \left[\mathrm{E}\widehat{f}_n\left(x + \frac{v_1}{M_n^{(2)}}, y - \frac{v_2}{L_n^{(2)}}\right) - f(x,y) \right] dv_1 dv_2.$$

$$E_9 = \frac{1}{\left(1-\frac{M_n^{(2)}}{M_n^{(1)}}\right)\left(1-\frac{L_n^{(2)}}{L_n^{(1)}}\right)} \int_{\frac{M_n^{(2)}}{M_n^{(1)}}}^{1} \int_{\frac{L_n^{(2)}}{L_n^{(1)}}}^{1} W^{(2)}(v_1) W^{(4)}(v_2) \times \left[\mathrm{E}\widehat{f}_n\left(x + \frac{v_1}{M_n^{(2)}}, y + \frac{v_2}{L_n^{(2)}}\right) - f(x,y) \right] dv_1 dv_2.$$

We group all expressions in the form noted E':

$$E' = \frac{1}{\left(1-\frac{M_n^{(2)}}{M_n^{(1)}}\right)\left(1-\frac{L_n^{(2)}}{L_n^{(1)}}\right)} \int_{\frac{M_n^{(2)}}{M_n^{(1)}}}^{1} \int_{\frac{L_n^{(2)}}{L_n^{(1)}}}^{1} W^{(2)}(v_1) W^{(4)}(v_2) \times \left[\mathrm{E}\widehat{f}_n\left(x \pm \frac{v_1}{M_n^{(2)}}, y \pm \frac{v_2}{L_n^{(2)}}\right) - f(x,y) \right] dv_1 dv_2. \tag{14}$$

Let us now show that E' tends to 0, indeed we can write

$$\mathrm{E}\widehat{f}_n\left(x \pm \frac{v_1}{M_n^{(2)}}, y \pm \frac{v_2}{L_n^{(2)}}\right) - f(x,y) = \mathrm{R}'_n + \mathrm{S}'_n + \mathrm{D}'_n \text{ where}$$

$$\mathrm{R}'_n = \frac{1}{h_n^2} \int K\left(\frac{x \pm \frac{v_1}{M_n^{(2)}} - z_1}{h_n}, \frac{y \pm \frac{v_2}{L_n^{(2)}} - z_2}{h_n} \right) f(z_1, z_2) dz_1 dz_2 - f(x,y)$$

$$\mathrm{S}'_n = \frac{1}{h_n^2}\sum_{j=1}^{q} a'_j\, K\left(\frac{x \pm \dfrac{v_1}{M_n^{(2)}} - w_{1j}}{h_n}, \frac{y \pm \dfrac{v_2}{L_n^{(2)}} - w_{2j}}{h_n}\right)$$

$$\mathrm{D}'_n = \frac{1}{h_n^2}\sum_{j=1}^{q'} \int K\left(\frac{x \pm \dfrac{v_1}{M_n^{(2)}} - u_1}{h_n}, \frac{y \pm \dfrac{v_2}{L_n^{(2)}} - a_i u_1 - b_i}{h_n}\right)\varphi_i(u_1)du_1 \,.$$

Since f is uniformly continuous, R_n tends to zero uniformly in v_1, v_2 of $[-1,1]$. The rate of convergence is $O(h_n^2)$ see (Bosq and Lecoutre 1987).

$$\mathrm{S}'_n = \frac{1}{h_n^2}\sum_{j=1}^{q} a'_j\, K_1\left(\frac{x \pm \dfrac{v_1}{M_n^{(2)}} - w_{1j}}{h_n}\right) K_2\left(\frac{y \pm \dfrac{v_2}{L_n^{(2)}} - w_{2j}}{h_n}\right)$$

We distinguish the four following cases:

1) If $(x, y) = (w_{1j}, w_{2j})$

Since $\dfrac{1}{M_n^{(1)}} \le \dfrac{v_1}{M_n^{(2)}} \le \dfrac{1}{M_n^{(2)}}$ and $\dfrac{1}{L_n^{(1)}} \le \dfrac{v_2}{L_n^{(2)}} \le \dfrac{1}{L_n^{(2)}}$. As K_1 and K_2 are decreasing functions, we obtain

$$\mathrm{S}'_n = O\left(\frac{1}{h_n^2} K_1\left(\frac{1}{h_n M_n^{(1)}}\right) K_2\left(\frac{1}{L_n^{(1)} h_n}\right)\right).$$

2) If $x \ne w_{1j}$ and $y \ne w_{2j}$, it is clear that $\mathrm{S}'_n = O\left(\dfrac{1}{h_n^2} K_1\left(\dfrac{1}{h_n}\right) K_2\left(\dfrac{1}{h_n}\right)\right)$.

3) If $x = w_{1j}$ and $y \ne w_{2j}$, it is clear that $\mathrm{S}'_n = O\left(\dfrac{1}{h_n^2} K_1\left(\dfrac{1}{M_n^{(1)} h_n}\right) K_2\left(\dfrac{1}{h_n}\right)\right)$.

4) If $x \ne w_{1j}$ and $y = w_{2j}$, we get $\mathrm{S}'_n = O\left(\dfrac{1}{h_n^2} K_1\left(\dfrac{1}{h_n}\right) K_2\left(\dfrac{1}{L_n^{(1)} h_n}\right)\right)$.

On the other hand, the expression D'_n can be written as the sum of D_{ni} where D_{ni} is defined as

$$D_{ni} = \frac{1}{h_n^2} \int K\left(\frac{x \pm \frac{v_1}{M_n^{(2)}} - u_1}{h_n}, \frac{y \pm \frac{v_2}{L_n^{(2)}} - a_i u_1 - b_i}{h_n} \right) \varphi_i(u_1) du_1 .$$

a) If $y \neq a_i x + b_i$ from the equality (10) we get:

$$D_{ni} = O\left(\frac{1}{h_n} K_1\left(\frac{1}{h_n} \right) \right) + O\left(\frac{1}{h_n} K_2\left(\frac{1}{h_n} \right) \right).$$

b) If $y = a_i x + b_i$ we have

$$D_{ni} = \frac{1}{h_n^2} \int K_1\left(\frac{x \pm \frac{v_1}{M_n^{(2)}} - u_1}{h_n} \right) K_2\left(\frac{\pm \frac{v_2}{L_n^{(2)}} + a_i\left(x - u_1\right)}{h_n} \right) \varphi_i(u_1) du_1$$

$$= \frac{1}{h_n^2} \int K_1\left(\frac{t \pm \frac{v_1}{M_n^{(2)}}}{h_n} \right) K_2\left(\frac{\pm \frac{v_2}{L_n^{(2)}} + a_i t}{h_n} \right) \varphi_i(x - t) dt$$

– If $a_i = 0$, we have $D_{ni} = O\left(\frac{1}{h_n} K_2\left(\frac{1}{L_n^{(1)} h_n} \right) \right)$.

– Let us show that for $a_i \neq 0$ when n is large, the numerators of K_1 and K_2 are not vanishing at same value t. Indeed we assume that there is a real number t such that $t = \pm \frac{v_1}{M_n^{(2)}}$ and $t = \pm \frac{v_2}{a_i L_n^{(2)}}$, therefore $\frac{L_n^{(2)}}{M_n^{(2)}} = \pm \frac{v_2}{a_i v_1}$ Since $L_n^{(2)}$, $M_n^{(2)}$, v_2 and v_1 are nonnegative, the last equality becomes $\frac{L_n^{(2)}}{M_n^{(2)}} = \frac{v_2}{|a_i| v_1}$. We choose $M_n^{(2)}$ and $L_n^{(1)}$ such that $M_n^{(2)} > a_i L_n^{(1)}$. Since $\frac{L_n^{(2)}}{L_n^{(1)}} < v_2 < 1$ and $\frac{M_n^{(2)}}{M_n^{(1)}} < v_1 < 1$, we get that $\frac{L_n^{(2)}}{a_i L_n^{(1)}} < \frac{v_2}{a_i v_1} < \frac{M_n^{(1)}}{M_n^{(2)}}$ and therefore $\frac{L_n^{(2)}}{a_i L_n^{(1)}} < \frac{L_n^{(2)}}{M_n^{(2)}}$ contradicts with the fact that $M_n^{(2)} > a_i L_n^{(1)}$.

Without losing the generality, we assume that $\pm\frac{v_1}{M_n^{(2)}} < \pm\frac{v_2}{a_i L_n^{(2)}}$. We can split the integral of D_{ni} as follows:

$$D_{ni} = \frac{1}{h_n^2}\int_{-\infty}^{\pm\frac{v_1}{M_n^{(2)}}-\eta} K_1\left(\frac{\pm\frac{v_1}{M_n^{(2)}}-t}{h_n}\right)K_2\left(\frac{\pm\frac{v_2}{L_n^{(2)}}+a_i t}{h_n}\right)\varphi_i(t+x)dt$$

$$+\frac{1}{h_n^2}\int_{\pm\frac{v_1}{M_n^{(2)}}-\eta}^{\pm\frac{v_1}{M_n^{(2)}}+\eta} K_1\left(\frac{\pm\frac{v_1}{M_n^{(2)}}-t}{h_n}\right)K_2\left(\frac{\pm\frac{v_2}{L_n^{(2)}}+a_i t}{h_n}\right)\varphi_i(t+x)dt$$

$$+\frac{1}{h_n^2}\int_{\pm\frac{v_1}{M_n^{(2)}}+\eta}^{\pm\frac{v_2}{a_i L_n^{(2)}}-\eta} K_1\left(\frac{\pm\frac{v_1}{M_n^{(2)}}-t}{h_n}\right)K_2\left(\frac{\pm\frac{v_2}{L_n^{(2)}}+a_i t}{h_n}\right)\varphi_i(t+x)dt$$

$$+\frac{1}{h_n^2}\int_{\pm\frac{v_2}{a_i L_n^{(2)}}-\eta}^{\pm\frac{v_2}{a_i L_n^{(2)}}+\eta} K_1\left(\frac{\pm\frac{v_1}{M_n^{(2)}}-t}{h_n}\right)K_2\left(\frac{\pm\frac{v_2}{L_n^{(2)}}+a_i t}{h_n}\right)\varphi_i(t+x)dt$$

$$+\frac{1}{h_n^2}\int_{\pm\frac{v_2}{a_i L_n^{(2)}}+\eta}^{+\infty} K_1\left(\frac{\pm\frac{v_1}{M_n^{(2)}}-t}{h_n}\right)K_2\left(\frac{\pm\frac{v_2}{L_n^{(2)}}+a_i t}{h_n}\right)\varphi_i(t+x)dt$$

$$D_{ni} = Z_1 + Z_2 + Z_3 + Z_4 + Z_5$$

where η is a nonnegative real number satisfying $\pm\frac{v_1}{M_n^{(2)}}-\eta < \pm\frac{v_2}{a_i L_n^{(2)}}-\eta$. In the first integral the numerators of K_1 and K_2, are not vanishing. K_1 and K_2 are continuous, decreasing, even functions. Thus Z_1 is bounded by $\left(\frac{1}{h_n^2}\right)\sup_{t\in\left]-\infty,\pm\frac{v_1}{M_n^{(2)}}-\eta\right[} K_1\left(\frac{\pm\frac{v_1}{M_n^{(2)}}-t}{h_n}\right)K_2\left(\frac{\pm\frac{v_2}{L_n^{(2)}}+a_i t}{h_n}\right)\int_{-\infty}^{\pm\frac{v_1}{M_n^{(2)}}-\eta}\varphi_i(t+x)dt$ which is

$O\left(\frac{1}{h_n^2}K_1\left(\frac{1}{h_n}\right)K_2\left(\frac{1}{h_n}\right)\right)$. In the same way we show that Z_3 and Z_5 have the same rate: $O\left(\frac{1}{h_n^2}K_1\left(\frac{1}{h_n}\right)K_2\left(\frac{1}{h_n}\right)\right)$.

In the second integral, since the numerator of K_2 is not vanishing, we increase K_2 as follows:

$$Z_2 \leq \left(\frac{1}{h_n^2}\right) \sup_{t\in\left[\pm\frac{v_1}{M_n^{(2)}}-\eta,\pm\frac{v_1}{M_n^{(2)}}+\eta\right]} K_2\left(\frac{\pm\frac{v_2}{L_n^{(2)}}+a_i t}{h_n}\right) \int_{\pm\frac{v_1}{M_n^{(2)}}-\eta}^{\pm\frac{v_1}{M_n^{(2)}}+\eta} K_1\left(\frac{\pm\frac{v_1}{M_n^{(2)}}-t}{h_n}\right)\varphi_i(t+x)dt.$$

As K_1 is a kernel, φ_i is uniformly continuous, we obtain $Z_2 = O\left(\frac{1}{h_n}K_2\left(\frac{1}{h_n}\right)\right)$. In the same manner it is shown that: $Z_4 = O\left(\frac{1}{h_n}K_1\left(\frac{1}{h_n}\right)\right)$. Thus we obtain

$$D_n = O\left(\frac{1}{h_n}K_1\left(\frac{1}{h_n}\right)\right) + O\left(\frac{1}{h_n}K_2\left(\frac{1}{h_n}\right)\right). \tag{15}$$

4.3. *Proof of the theorem 2.3*

By using the same arguments used to show the equality (14); it is easily shown that:

$$\mathrm{Var}(\widehat{g_n}(x,y)) = Z_n \mathrm{E}\left(\sum_{k,k'=1}^{2} J\left(k,k'\right) - \mathrm{E}(J(k,k'))\right)^2,$$

where $Z_n = \dfrac{1}{\left(1-\dfrac{M_n^{(2)}}{M_n^{(1)}}\right)^2}\dfrac{1}{\left(1-\dfrac{L_n^{(2)}}{L_n^{(1)}}\right)^2}$ and

$$J(k,k') = \int_{\frac{M_n^{(2)}}{M_n^{(1)}}}^{1}\int_{\frac{L_n^{(2)}}{L_n^{(1)}}}^{1} W^{(2)}(v_1)W^{(4)}(v_2)\widehat{f_n}\left(x+(-1)^k\frac{v_1}{M_n^{(2)}}, y+(-1)^{k'}\frac{v_2}{L_n^{(2)}}\right)dv_1 dv_2.$$

Therefore

$$\operatorname{Var}(\widehat{g_n}(x,y)) = Z_n \sum_{k,k',p,p'=1}^{2} \int_{d_n} W^{(2)}(v_1)W^{(4)}(v_2)W^{(2)}(v'_1)W^{(4)}(v'_2)C(v)dv$$

where $v=(v_1,v_2,v'_1,v'_2)$, $d_n = \left(\left[\frac{M_n^{(2)}}{M_n^{(1)}},1\right]\times\left[\frac{L_n^{(2)}}{L_n^{(1)}},1\right]\right)^2$ and

$$C(v) = \operatorname{Cov}\left(\widehat{f_n}\left[x+(-1)^k\frac{v_1}{M_n^{(2)}}, y+(-1)^{k'}\frac{v_2}{L_n^{(2)}}\right], \widehat{f_n}\left[x+(-1)^p\frac{v'_1}{M_n^{(2)}}, y+(-1)^{p'}\frac{v'_2}{L_n^{(2)}}\right]\right).$$

From Cauchy-Schwartz inequality, we have

$$C(v) \le \left(\operatorname{Var}\widehat{f_n}\left(x+(-1)^k\frac{v_1}{M_n^{(2)}}, y+(-1)^{k'}\frac{v_2}{L_n^{(2)}}\right)\right)^{\frac{1}{2}} \times \left(\operatorname{Var}\widehat{f_n}\left(x+(-1)^p\frac{v'_1}{M_n^{(2)}}, y+(-1)^{p'}\frac{v'_2}{L_n^{(2)}}\right)\right)^{\frac{1}{2}}.$$

It is easy to show that

$$\begin{aligned}\operatorname{Var}\widehat{f_n}\left(x\pm\frac{v_1}{M_n^{(2)}}, y\pm\frac{v_2}{L_n^{(2)}}\right) &= \frac{1}{n^2h_n^4}\sum_{i=1}^{n}\mathrm{E}\left(K^2\left(\frac{x\pm\frac{v_1}{M_n^{(2)}}-x_i}{h_n}, \frac{y\pm\frac{v_2}{L_n^{(2)}}-y_i}{h_n}\right)\right) \\ &\quad -\frac{1}{n}\mathrm{E}^2\left(\widehat{f_n}\left(x\pm\frac{v_1}{M_n^{(2)}}, y\pm\frac{v_2}{L_n^{(2)}}\right)\right) \qquad (16) \\ &= H_1+H_2+H_3-\frac{1}{n}\mathrm{E}^2\left(\widehat{f_n}\left(x\pm\frac{v_1}{M_n^{(2)}}, y\pm\frac{v_2}{L_n^{(2)}}\right)\right)\end{aligned}$$

where H_i are defined by:

$$H_1 = \frac{1}{nh_n^4}\int K^2\left(\frac{x\pm\frac{v_1}{M_n^{(2)}}-z_1}{h_n}, \frac{y\pm\frac{v_2}{L_n^{(2)}}-z_2}{h_n}\right)f(z_1,z_2)dz_1dz_2$$

$$H_2 = \frac{1}{nh_n^4}\sum_{j=1}^{q} a_j K^2\left(\frac{x \pm \frac{v_1}{M_n^{(2)}} - w_{1j}}{h_n}, \frac{y \pm \frac{v_2}{L_n^{(2)}} - w_{2j}}{h_n}\right)$$

$$H_3 = \frac{1}{nh_n^4}\int\sum_{i=1}^{q'} K^2\left(\frac{x \pm \frac{v_1}{M_n^{(2)}} - u}{h_n}, \frac{y \pm \frac{v_2}{L_n^{(2)}} - a_i u - b_i}{h_n}\right)\varphi_i(u)du.$$

Putting $x \pm \frac{v_1}{M_n^{(2)}} - z_1 = h_n t_1$ and $y \pm \frac{v_2}{L_n^{(2)}} - z_2 = h_n t_2$, in the integral of H_1, we obtain

$$H_1 = \frac{1}{nh_n^2}\int K^2(t_1,t_2) f\left(t_1 h_n - x \pm \frac{v_1}{M_n^{(2)}}, t_2 h_n - y \pm \frac{v_2}{L_n^{(2)}}\right) dt_1 dt_2 .$$

From the theorem of the finite increments, we have

$$H_1 \le \frac{1}{nh_n^2}\int K^2(t_1,t_2)\max(f')\|(t_1h_n, t_2h_n)\| dt_1 dt_2 + \frac{1}{nh_n^2}\int K^2(t_1,t_2) f\left(\pm\frac{v_1}{M_n^{(2)}}, y \pm \frac{v_2}{L_n^{(2)}}\right) dt_1 dt_2 .$$

Thus we obtain $H_1 = O\left(\frac{1}{nh_n}\right) + O\left(\frac{f(x,y)}{nh_n^2}\iint K^2\right)$.

For H_2 we distinguish four cases:

a) If $(x,y) = (w_{1j}, w_{2j})$, we obtain $H_2 = O\left(\frac{1}{nh_n^4}K^2\left(\frac{1}{M_n^{(1)}h_n}, \frac{1}{L_n^{(1)}h_n}\right)\right)$.

b) If $x = w_{1j}$ and $y \ne w_{2j}$, it is easy to show that $H_2 = O\left(\frac{1}{nh_n^4}K^2\left(\frac{1}{M_n^{(1)}h_n}, \frac{1}{h_n}\right)\right)$.

c) If $x \ne w_{1j}$ and $y = w_{2j}$, we easily obtain $H_2 = O\left(\frac{1}{nh_n^4}K^2\left(\frac{1}{h_n}, \frac{1}{L_n^{(1)}h_n}\right)\right)$.

d) If $x \ne w_{1j}$ and $y \ne w_{2j}$, it is clear that $H_2 = O\left(\frac{1}{nh_n^4}K^2\left(\frac{1}{h_n}, \frac{1}{h_n}\right)\right)$.

On the other hand, we note $H_3 = \sum_{i=1}^{q'} G_i$ where

$$G_i = \frac{1}{nh_n^4} \int K_1^2 \left(\frac{x \pm \frac{v_1}{M_n^{(2)}} - u}{h_n} \right) K_2^2 \left(\frac{y \pm \frac{v_2}{M_n^{(2)}} - a_i u - b_i}{h_n} \right) \varphi_i(u) du.$$

Distinguish the two following cases:

1) If $y \neq a_i x + b_i$ from (10), we have

$$G_i = O\left(\frac{1}{nh_n^3} K_1^2 \left(\frac{1}{h_n} \right) \right) + O\left(\frac{1}{nh_n^3} K_2^2 \left(\frac{1}{h_n} \right) \right).$$

2) If $y = a_i x + b_i$, we get the following:

$$G_i = \frac{1}{nh_n^4} \int K_1^2 \left(\frac{x \pm \frac{v_1}{M_n^{(2)}} - u}{h_n} \right) K_2^2 \left(\frac{\pm \frac{v_2}{M_n^{(2)}} - a_i (x - u)}{h_n} \right) \varphi_i(u) du.$$

By a similar work used to show (15), namely splitting the integral to 5 integrals under the neighbourhoods of the points where the numerators of K_1 and K_2 are respectively vanishing and the remaining points, we get

$$G_i = O\left(\frac{1}{nh_n^3} K_1^2 \left(\frac{1}{h_n} \right) \right) + O\left(\frac{1}{nh_n^3} K_2^2 \left(\frac{1}{h_n} \right) \right).$$

Thus we conclude the result of this theorem.

4.4. *Proof of the theorem 3.1*

From (1) we have

$$\begin{aligned} \mathrm{E}(\widehat{a_n}(x,y)) &= \frac{1}{K(0,0)} \iint K\left(\frac{x - z_1}{h_n}, \frac{y - z_2}{h_n} \right) f(z_1, z_2) dz_1 dz_2 \\ &+ \frac{1}{K(0,0)} \sum_{j=1}^{q} a'_j K\left(\frac{x - w_{1j}}{h_n}, \frac{y - w_{2j}}{h_n} \right) \\ &+ \frac{1}{K(0,0)} \sum_{i=1}^{q'} \int K\left(\frac{x - v_1}{h_n}, \frac{y - a_i v_1 - b_i}{h_n} \right) \varphi_i(v_1) dv_1 \\ &= A'_n + B'_n + C'_n \end{aligned}$$

As in the proof of theorem 2.1, we have

$$A'_n = \begin{cases} \dfrac{h_n^2}{K(0,0)} f(x,y) + O\left(\dfrac{h_n^4}{K(0,0)}\right) & \text{if} \quad (x,y) = (w_{1j}, w_{2j}) \\ O(h_n^2) & \text{if} \quad (x,y) \neq (w_{1j}, w_{2j}). \end{cases}$$

On the other hand, if $(x,y) \neq (w_{1j}, w_{2j})$ we have from (8),

$$B'_n = O\left(K_1\left(\frac{1}{h_n}\right) K_2\left(\frac{1}{h_n}\right)\right).$$

If $(x,y) = (w_{1j}, w_{2j})$ we obtain

$$B'_n = a'_j + \sum_{k \neq j}^{n} \left(K_1\left(\frac{w_{1j} - w_{1k}}{h_n}\right) K_2\left(\frac{w_{2j} - w_{2k}}{h_n}\right)\right).$$

$$B'_n = a'_j + O\left(K_1\left(\frac{1}{h_n}\right) K_2\left(\frac{1}{h_n}\right)\right).$$

For the term C'_n we distinguish two cases: if $(x,y) = (w_{1j}, w_{2j})$, therefore $y \neq a_i x + b_i$ for all $i = 1,2,...,q'$, from (10) we have

$$C'_n = O\left(h_n K_1\left(\frac{1}{h_n}\right)\right) + O\left(h_n K_2\left(\frac{1}{h_n}\right)\right).$$

If $y = a_i x + b_i$, we obtain

$$C'_n \leq \frac{1}{K(0,0)} h_n \sup(K_2) \sum_{i=1}^{q'} \int \frac{1}{h_n} K_1\left(\frac{x - v_1}{h_n}\right) \varphi_i(v_1) dv_1$$

Thus

$$C'_n = O(h_n).$$

Thus the result of this theorem follows.

For showing that the variance tends to zero, we use (16) we obtain

$$\mathrm{Var}(\widehat{a_n}(x,y)) = \frac{h_n^4}{K^2(0,0)} \mathrm{Var}\left(\frac{1}{nh_n^2} K\left(\frac{x-x_i}{h_n}, \frac{y-y_i}{h_n}\right)\right)$$
$$= \frac{h_n^4}{K^2(0,0)}\left(H_1 + H_2 + H_3 - \frac{1}{n}\mathrm{E}^2\,\widehat{f_n}(x,y)\right)$$

where

$$H_1 = \frac{1}{nh_n^2}\int K^2\left(\frac{x-z_1}{h_n}, \frac{y-z_2}{h_n}\right) f(z_1,z_2)dz_1dz_2.$$

$$H_2 = \frac{1}{nh_n^4}\sum_{j=1}^{q} a_j K^2\left(\frac{x-w_{1j}}{h_n}, \frac{y-w_{2j}}{h_n}\right),$$

$$H_3 = \frac{1}{nh_n^4}\int\sum_{j=1}^{q'} K^2\left(\frac{x-u}{h_n}, \frac{y-a_i u - b_i}{h_n}\right)\varphi_i(u)du.$$

It is easy to see that

$$\frac{h_n^4}{K^2(0,0)} H_1 = O\left(\frac{h_n^3}{n}\right) + O\left(\frac{h_n^2}{n}\right) f(x,y)\iint K^2$$

$$\frac{h_n^4}{K^2(0,0)} H_2 = \begin{cases} O\left(\frac{h_n}{n} K_1^2\left(\frac{1}{h_n}\right)\right) + O\left(\frac{h_n}{n} K_2^2\left(\frac{1}{h_n}\right)\right) & \text{if } (x,y) = (w_{1j}, w_{2j}) \\ O\left(\frac{1}{n}\right) & \text{if } (x,y) \neq (w_{1j}, w_{2j}) \end{cases}$$

Since K_1 and K_2 are bounded functions, we have

$$\frac{h_n^4}{K^2(0,0)} H_3 = \begin{cases} O\left(\frac{1}{n} K_1^2\left(\frac{1}{h_n}\right) K_2^2\left(\frac{1}{h_n}\right)\right) & \text{if } (x,y) \neq (w_{1j}, w_{2j}) \\ O\left(\frac{1}{n}\right) + O\left(\frac{1}{n} K_1^2\left(\frac{1}{h_n}\right) K_2^2\left(\frac{1}{h_n}\right)\right) & \text{if } (x,y) = (w_{1j}, w_{2j}) \end{cases}$$

Thus the result follows.

4.5. *Proof of the theorem 3.2*

Let λ_1 a real number and λ_2 is such $\lambda_2 = a_i\lambda_1 + b_i$ where $i \in 1,2,..q'$.

$$\widehat{\varphi_i}(\lambda_1) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_n}K\left(\frac{\lambda_1 - x_i}{h_n}, \frac{\lambda_2 - y_i}{h'_n}\right)$$

$$\begin{aligned}
\mathrm{E}(\widehat{\varphi_i}(\lambda_1)) &= \frac{1}{K_2(0)}\iint \frac{1}{h_n}K_1\left(\frac{\lambda_1 - z_1}{h_n}\right)K_2\left(\frac{\lambda_2 - z_2}{h'_n}\right)f(z_1, z_2)dz_1dz_2 \\
&+ \frac{1}{K_2(0)}\sum_{j=1}^{q}\frac{1}{h_n}a'_j\,K\left(\frac{\lambda_1 - w_{1j}}{h_n}, \frac{\lambda_2 - w_{2j}}{h'_n}\right) \\
&+ \frac{1}{K_2(0)}\sum_{i=1}^{q'}\int\frac{1}{h_n}K\left(\frac{\lambda_1 - v_1}{h_n}, \frac{\lambda_2 - a_i v_1 - b_i}{h'_n}\right)\varphi_i(v_1)dv_1 \\
&= A''_n + B''_n + C''_n
\end{aligned}$$

It is easy to see

$$A''_n = O(h_n)$$

$$B''_n = O\left(\frac{1}{h_n}K_1\left(\frac{1}{h_n}\right)K_2\left(\frac{1}{h'_n}\right)\right)$$

$$C''_n = \frac{1}{K_2(0)}\sum_{i=1}^{q'}\int\frac{1}{h_n}K\left(\frac{\lambda_1 - v_1}{h_n}, \frac{a_i(\lambda_1 - v_1)}{h'_n}\right)\varphi_i(v_1)dv_1$$

Putting $\frac{\lambda_1 - v_1}{h_n} = v$, we have

$$C''_n = \frac{1}{K_2(0)}\sum_{i=1}^{q'}\int K_1(v)K_2\left(\frac{a_i v h_n}{h'_n}\right)\varphi_i(\lambda_1 - v h_n)dv$$

Using the theorem of dominated convergence the result of theorem follows.

Let us show that the estimate is consistent

$$\mathrm{Var}\widehat{\varphi_i}(\lambda_1) = \frac{1}{K_2^2(0)}\left(H''_1 + H''_2 + H''_3 - \frac{1}{n}\mathrm{E}^2(\varphi_i(\lambda_1))\right)$$

where

$$H''_1 = \frac{1}{n h_n^2 h'^2_n}\int K^2\left(\frac{\lambda_1 - z_1}{h_n}, \frac{\lambda_2 - z_2}{h'_n}\right)f(z_1, z_2)dz_1dz_2$$

$$H''_2 = \frac{1}{nh_n^2 h'^2_n} \sum_{j=1}^{q} a'_j K^2\left(\frac{\lambda_1 - w_{1j}}{h_n}, \frac{\lambda_2 - w_{2j}}{h'_n}\right)$$

$$H''_3 = \frac{1}{nh_n^2 h'^2_n} \sum_{i=1}^{q'} \int K^2\left(\frac{\lambda_1 - u}{h_n}, \frac{a_i(\lambda_1 - u)}{h'_n}\right)\varphi_i(u)du.$$

Putting $\lambda_1 - z_1 = h_n t_1$ and $\lambda_2 - z_2 = h'_n t_2$ in the integral of H''_1 we obtain

$$H''_1 = \frac{1}{nh_n h'_n} \int K^2(t_1, t_2) f(t_1 h_n - x, t_2 h'_n - y) dt_1 dt_2$$

From the theorem of the finite increments, we have

$$H''_1 \leq \frac{1}{nh_n h'_n} \int K^2(t_1, t_2) \max(f') \left\|(t_1 h_n, t_2 h'_n)\right\| dt_1 dt_2$$

$$+ \frac{1}{nh_n h'_n} \int K^2(t_1, t_2) f(x, y) dt_1 dt_2.$$

Thus, as $\frac{h_n}{h'_n} \to 0$, we obtain $H''_1 = O\left(\frac{1}{nh_n}\right) + O\left(\frac{f(x,y)}{nh_n h'_n} \int\int K^2\right)$. On the other hand as $\lambda_1 \neq w_{1j}$ and $\lambda_2 \neq w_{2j}$, we have $H''_2 = O\left(\frac{1}{nh_n^2 h'^2_n} K^2\left(\frac{1}{h_n}, \frac{1}{h'_n}\right)\right)$.

Putting $\frac{\lambda_1 - u}{h_n} = v$, we write

$$H''_3 = \frac{1}{h_n^2 h'_n} \int K_1^2(v) K_2^2\left(a_i v \frac{h_n}{h'_n}\right) \varphi_i(\lambda_1 - v h_n) dv$$

Therefore

$$H''_3 = O\left(\frac{1}{nh_n^2 h'_n}\right).$$

Thus the result follows.

5. SIMULATION

We consider the univariate case. Let X be a random variable with measure:

$$du = f(x)dx + a_1 \delta_{\lambda_1} + a_2 \delta_{\lambda_2}$$

For the simulation we take f the density of the stand Gaussian random variable and we choose $\lambda_1 = -2$ and $\lambda_2 = 2$, this choice is arbitrary.

For having a sample $(x_1, x_2, ..., x_{1000})$ of the random variable X, we generate a sample $(u_1, u_2, ..., u_{1000})$ uniformly distributed over the interval $(0,1)$ and a sample $(y_1, y_2, ..., y_{1000})$ of the gaussian variable and for all $1 \le i \le 1000$ we test if $b_1 < u_i < b_2$ we take $x_i = \lambda_1$ and if $b_2 < u_i < b_3$, we take $x_i = \lambda_2$ otherwise we take $x_i = y_i$ with $b_1 = 0.5$, $b_2 = 0.55$ and $b_3 = 0.6$. The choice of the values of the parameters b_1, b_2 and b_3 influence only on the amplitudes a_1 and a_2. Since we take $b_2 - b_1 = b_3 - b_1$, the amplitudes have approximatively identical values.

We calculate the estimator, $\widehat{f_n}(x)$, defined in the section 2.3, with $K(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\}$, $h_n = n^{-1/5}$ and $n = 1000$. Repeating this simulation with same parameters for 20 samples, we remark that the graphics of $\widehat{f_n}(x)$ represent two jump points localized respectively in $[-2.3, -1.7]$ and $[1.7, 2.3]$. Thus we take $[\alpha_1, \beta_1] = [-2.3, -1.7]$ and $[\alpha_2, \beta_2] = [1.7, 2.3]$.

Using these samples, we calculate, for fixed x, the empirical mean of $\widehat{a'}_n(x) = \frac{h_n}{K(0)} \sum_{i=1}^{n} \frac{1}{nh_n} K\left(\frac{x - x_i}{h_n}\right)$ denoted by $\overline{a'}(x)$. For $x = \pm 2$, $\overline{a'}(x) \simeq 0.12$ and for all other points in the intervals $[\alpha_1, \beta_1]$ and $[\alpha_2, \beta_2]$ the value of $\overline{a'}$ is less than 0.001. Since f is the density of the stand Gaussian variable, we verify that, for $x = \pm 2$, $\widehat{f_n}(x) \simeq \overline{a'}(x) + f(x)$.

The bandwidth of spectral windows are taken as: $M_n^{(1)} = 8n^{1/6}$, $M_n^{(2)} = n^{1/7}$. The spectral windows are chosen as:

$$W^{(1)}(t) = \begin{cases} \frac{64}{63} t + \frac{64}{63} & \text{if } \; t \in [-1, -1/8] \\ 1/8 & \text{if } \; t \in [-1/8, 1/8] \\ -\frac{64}{63} t + \frac{64}{63} & \text{if } \; t \in [1/8, 1] \\ 0 & \text{otherwise} \end{cases}$$

$$W^{(2)}(t) = \begin{cases} \frac{31}{56} & \text{if } \; t \in [-1, -1/8] \cup [1/8, 1] \\ 1/8 & \text{if } \; t \in [-1/8, 1/8] \\ 0 & \text{otherwise} \end{cases}$$

The following graphics represent the estimate $\widehat{g_n}(x)$ defined in the section 2.3 and the density kernel of the variable X .

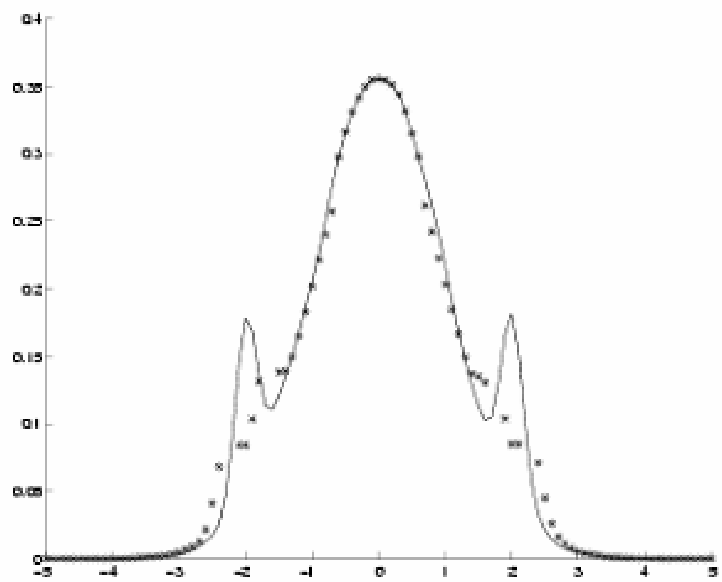

Figure 1 – Graphics of the estimate and kernel density of variable X: Asteriksed line is the estimate proposed and continuous line is the kernel density of X.

Université de Bourgogne à Dijon RACHID SABRE
France

ACKNOWLEDGMENTS

I would like to thank the referees for their interest in this paper and their valuable comments and suggestions.

REFERENCES

L. BIGGERI, (1999), *Diritto alla 'privacy' e diritto all'informazione statistica*, in Sistan-Istat, "Atti della Quarta Conferenza Nazionale di Statistica", Roma, 11-13 novembre 1998, Roma, Istat, Tomo 1, pp. 259-279.

M. BERTRAND-RETALI, (1978), *Convergence d'un estimateur de la densité par la méthode du noyau,* Rev. Roum. de Mathématiques pures et Appl. T XXIIII, N 3, 361-385.

P. BILLINGSLEY , (1968), *Convergence of probability measures*, Ed. Willey, New York.

D. BOSQ , J.P. LECOUTRE, (1987), *Théorie de l'estimation fonctionnelle*, Ed. Economica, Paris.

A. W. BOWMAN, P. HALL AND D. M. TITTRINGTON, (1984), *Cross-validation in nonparametric estimation of probabilities and probability density*, Biometrika, 71, 341-352.

P. DEHEUVELS, (1977), *Estimation non paramétrique de la densité par histogrammes généralisés* , Revue de Statistique Appliquée, 25, f.3, 5-42.

P. DEHEUVELS, P HOMINAL, (1979), *Estimation non paramétrique de la densité compte tenu d'informations sur le support* , Revue de Statistique Appliquée, 27, f.3, 47-68.

P. DEHEUVELS, P HOMINAL, (1980), *Estimation automatique de la densité,* Revue de Statistique Appliquée, 28, f.1, 23-55.

L. DEVROYE, (1987), *A course in density estimation*, Ed. Birkhauser, Boston INC.

J. FAN, (1991) *On the optimal rates of convergence for nonparametric deconvolution problems*, Annals of Statistics, 19, 1257-1272.

E. PARZEN, (1962), *On the estimation of a probability density function and mode*, Annals of Mathematical Statistics, 33, 1065-1076.

M. ROSEMBLATT, (1965), *Remarks on some nonparametric estimates of a density function*, Annals of Mathematical Statistics, 27, 832-837.

M. RACHDI , R. SABRE, (2000) *Consistent estimates of the mode of the probability density function in nonparametric deconvolution*, Statistics & Probability Letters, 47(2000), 105-114. 65-78.

R. SABRE, (1994), *Estimation de la densit´e de la mesure spectrale mixte pour un processus symétrique stable strictement stationaire*, C. R. Acad. Sci. Paris, t. 319, Serie I, p. 1307-1310.

R. SABRE, (1995), *Spectral density estimation for stationary stable random fields*, Journal Applicationes Mathematicae, 23, 2, p. 107-133.

A. STEFANSKI, (1990), *Rates of convergence of some estimators in a class of deconvolution problems*, Statistics & Probability Letters, 9, p. 229-235.

P. VIEU, (1996), *A note on density mode estimation*, Statistics & Probability Letters, 26, p. 297-307.

C.H. ZHANG, (1990), *Fourier methods for estimating mixing densities and distributions*, Annals of Statistics, 18, p. 806-830.

RIASSUNTO

Stima di densità nonparametrica della parte continua di una misura mista

Nel lavoro si considera una coppia di variabili aleatorie (X, Y) la cui misura di probabilità è la somma di una misura assolutamente continua, una misura discreta e un numero finito di misure assolutamente continue su diverse rette. Viene proposta una stima consistente e asintoticamente corretta della densità della parte continua e se ne determina la velocità di convergenza.

SUMMARY

Nonparametric density estimation of continuous part of a mixed measure

We consider a pair of random variables (X, Y) whose probability measure is the sum of an absolutely continuous measure, a discrete measure and a finite number of absolutely continuous measures on several lines (1). An asymptotically unbiased and consistent estimate, at all points, of the density of the continuous part is given as well as its rate of convergence. We also estimate the amplitude of the discrete measure and the densities on several lines.