Deconvolution Estimation of the Density for Mixed Measure

Filali Mostafa*

IMB, CNRS UMR 5584

Université de Bourgogne, 9 avenue Savary

F-21078 Dijon, France

*Corresponding author

Sabre Rachid

LE2I, CNRS, UMR 6306

Université de Bourgogne, Agrosup

26 Av Docteur Petitjean 21000 Dijon, France

Abstract—**Consider a random variable whose measure probability is the sum of an absolutely continuous part with respect to the Lebesgue measure and a discrete measure. In this work is provided a deconvolution estimator density of the continuous part and also the estimated amplitudes of discrete part when the observations ara taken with a noise error.**

Keywords-component; density probability; kernel estimate; deconvolution

I. INTRODUCTION

This paper consider a random variable X whose probability measure μ is the sum of an absolutely continuous with respect to the Lebesgue measure and a discrete measure:

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

where the number q is supposed nonnegative integer 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_m is the amplitude of the jump at λ_m and is assumed unknown. δ is the Dirac measure. The jump points (λ_m are known real numbers. The estimation of the continuous part f is given in general case see sabre[11]. A statistical test used to verify the existence of jump at any point is studied in sabre[12].

This work considers the case where the observation is taken with errors: $Y_i = X_i + \varepsilon_i$,

where the errors ε_i are independent and identically distributed (i.i.d) random variables with known density h, and the random variables ε_i and X_i are independent. Such a model of measurement contaminated with error exists in many different fields such as (physics, economics, medical, biology …), where the measurements cannot be observed directly. The practical problems of deconvolution can be found in Medgyessy [8] and Carroll and Ruppert [4], Van[14], Wang [15] .

Our goal is to estimate the density $f(x)$ for all real numbers x and the amplitude a_m of the jump points. So we consider two cases: the first case, when x is outside of the jump point, $\left(x \notin \cup_{m=1}^{q} [\alpha_m, \beta_m] \right)$, we use the deconvolution kernel estimate as in Carroll and Hall [3], Stefanski and Carroll [13], Fan [1][2] and Masry [5], [6], [7]. For the other points $\left(x \in \cup_{m=1}^{q} [\alpha_m, \beta_m] \right)$, we smooth the deconvolution estimate by two windows satisfying some conditions as in Sabre [11] [10] [9]. This paper gives an estimator $\hat{a}_m(x)$ converging to the amplitude a_m if $x = \lambda_m$ or converging to zero if $x \neq \lambda_m$. We can use this result to verify the presence of the jump at any point.

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 observed data 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 $Y = g(X) + \varepsilon$, ε must be a centered Gaussian variable. To show that we take a sample of the residues $\varepsilon_i = Y_i - g(X_i)$ $i = 1, 2, ... n$. If the empirical residuals distribution is Gaussian except at discrete points, for example λ_1 and λ_2, so the law of ε can be written as:

$$d\mu = f(x)dx + a_1 \delta_{\lambda_1} + a_2 \delta_{\lambda_2}. \tag{2}$$

If we show that f is the density of a centered gaussian variable, we are sure that we must change our model by adding an other discrete variable.

A concrete application of this model concerns the measurement of the surface of a plant's submarine taken by

aerial images in different areas of the Atlantic Ocean. Theoretically, the images are taken with iid random errors already known by the device manufacturer of digital imaging, it is a centered normal distribution. Thus we can consider the model $Y_i = X_i + \varepsilon_i$. Moreover, the random passage of group of fish influences the distribution of data by increasing the observed to reach a constant value. Thus we consider that the measure is the sum of two measures, one continuous and the other discrete.

This paper is organized as follows. In the second section we estimate the density outside the neighborhood of the jump point and we study the estimation of the density inside the neighborhood of jump point. In the third section we estimate the amplitude of the jump point and we give the exact rates of convergence of the estimate. We finish by studying the numerical simulation of this estimator.

II. KERNEL ESTIMATE OF THE DENSITY FUNCTION

Consider a random variable X whose probability measure, μ, is defined by (1). Our goal is to estimate, for every real number x, the density function f. Let K be a bounded even decreasing kernel such that $\lim_{y \to \infty} | yK(y) |= 0$. Φ_h, Φ_X, Φ_Y and Φ_K are the characteristic functions of ε, X, Y and K respectively. The independence of X and ε gives: $\Phi_Y(t) = \Phi_X(t)\Phi_\varepsilon(t)$.

Suppose that the characteristics functions Φ_h and Φ_K satisfy the following hypothesis:

a) $| \Phi_h(t) |> 0$

b) $\Phi_K(t) / \Phi_h\left(\dfrac{t}{b}\right) \in L^1 \bigcap L^\infty$, for all real number b.

To estimate the density function $f(x)$ we distinguish two cases: when $x \in A = \bigcup_{m=1}^{q}[\alpha_m, \beta_m]$ and $x \notin A$.

A. The Estimation of the Function f Outside A

Let x be a real outside of the jump point $(x \notin A)$. As in Stefanski et Carroll [13] et Fan [2] we consider $\hat{f}_n(x)$ the deconvolution kernel estimate of $f(x)$:

$$\hat{f}_n(x) = \frac{1}{nb_n}\sum_{j=1}^{n} W_{b_n}\left(\frac{x-Y_j}{b_n}\right), \text{ where}$$

$$W_{b_n}(x) = \frac{1}{2\pi}\int_R e^{-itx}\frac{\Phi_K(t)}{\Phi_h(t/b_n)}dt, \tag{3}$$

and $\{b_n\}_{n=1}^{\infty}$ is a positive real sequence such that b_n converges to zero. From b) it is easy to see that W_b is integrable and $\int W_b(t)dt = 1$. The following results show that $\hat{f}_n(x)$ is asymptotically unbiased estimator of $f(x)$ and gives the rate of the convergence of the bias.

Lemme 1 *Let x be a real outside A, then $\hat{f}_n(x)$ is an asymptotically unbiased estimate. If f is twice differentiable and its partial derivatives are continuous and bounded and $\int x^2 K(x)dx < \infty$, we have:*

$$E\left(\hat{f}_n(x)\right) - f(x) = O\left(b_n^2\right) + O\left(\frac{1}{b_n}K\left(\frac{1}{b_n}\right)\right).$$

Proposition 1 *Under the conditions of lemma and chosen b_n such that $nb_n^{2\beta+1}$ converges to infinity, then*

$$var\left(\hat{f}_n(x)\right) = O\left(\frac{1}{nb_n^{2\beta+1}}\left(1+\sum_{m=1}^{q}a_m\right)D\right), \quad where$$

$$D = (1/2\pi B)^2 \int_{-\infty}^{+\infty}|t|^{2\beta}\left|\Phi_K(t)\right|^2 dt.$$

From Fan [2] , it is easy to show the lemma1 and proposition1

B. Estimation of f in Neighborhood of Jump Points

Let x belong to A. As in Sabre[11], we smooth the deconvolution estimator by two windows defined as follows:

$$\hat{g}_n(x) = \int S_n(x-u)\hat{f}_n(u)du,$$

where $\hat{f}_n$ is defined in (2) and $S_n(v)$ is defined by:

$$S_n(v) = \frac{U_n^{(2)}(v) - \dfrac{M_n^{(2)}}{M_n^{(1)}}U_n^{(1)}(v)}{1 - \dfrac{M_n^{(2)}}{M_n^{(1)}}},$$

(2)where

where the windows functions $U_n^{(1)}$ and $U_n^{(2)}$ satisfy:

$$U_n^{(1)}(t) = M_n^{(1)}U^{(1)}\left(tM_n^{(1)}\right); \quad U_n^{(2)}(t) = M_n^{(2)}U^{(2)}\left(tM_n^{(2)}\right)$$

with $M_n^{(1)}$ and $M_n^{(2)}$ are positive real sequences satisfying

$$\frac{M_n^{(2)}}{M_n^{(1)}} \to 0 \text{ and } M_n^{(r)} \to +\infty; \quad M_n^{(r)}b_n \to 0, \ r = 1, 2$$

The functions $U^{(1)}$ and $U^{(2)}$ are positive even integrable, vanishing outside of the interval $[-1,1]$ and satisfying $\int_{-1}^{1} U^{(i)}(x)\,dx = 1$, for $i = 1, 2$, and

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

Choose K such that $\dfrac{1}{b_n} K\left(\dfrac{1}{b_n M_n^{(1)}}\right)$ converges to zero.

Theorem 1 *Let x belonging to A.*

If f is twice differentiable and its partial derivatives are continuous and bounded, we have

$$E\left(\hat{g}_n(x)\right) - f(x) = O\left(\frac{1}{M_n^{(2)}}\right) + O\left(\frac{1}{b_n} K\left(\frac{1}{b_n M_n^{(1)}}\right)\right).$$

2) Under the conditions of lemma 1 and suppose that $nb_n^{2\beta+1}$ goes to infinity, then

$$var\left(\hat{g}_n(x)\right) = O\left(\frac{1}{nb_n^{2\beta+1}}\left(1 + \sum_{m=1}^{q} a_m\right) D\right),$$

Proof:

1) Using the definition of $\hat{g}_n$ in (14), we obtain

$$E\left(\hat{g}_n(x) - f(x)\right) = \int S_n(x-u)\left(E\hat{f}_n(u) - f(x)\right) du$$

$$= \int_{-\infty}^{x-\frac{1}{M_n^{(1)}}} + \int_{x-\frac{1}{M_n^{(1)}}}^{x+\frac{1}{M_n^{(1)}}} + \int_{x+\frac{1}{M_n^{(1)}}}^{+\infty}$$

$$\overset{\Delta}{=} E_1 + E_2 + E_3.$$

To show that E_2 is null, we use the following change of variable $x - u = v$.

Now showing that E_1 and E_3 converge to zero.

utting $x - u = v$ in the integral of E_3 we obtain

$$E_3 = \frac{1}{1 - \dfrac{M_n^{(2)}}{M_n^{(1)}}} \int_{-\infty}^{-\frac{1}{M_n^{(1)}}} D_n(v) P_n(v)\,dv, \text{ where}$$

$$D_n(v) = \left(M_n^{(2)} U^{(2)}\left(M_n^{(2)}v\right) - M_n^{(2)} U^{(1)}\left(M_n^{(1)}v\right)\right) \text{ and}$$

$$P_n(v) = \left(E\hat{f}_n(x-v) - f(x)\right)$$

The function $U^{(1)}$ being null outside of $[-1,1]$, then $U^{(1)}\left(M_n^{(1)}v\right)$ is null when v belongs to $\left[-\infty, -\dfrac{1}{M_n^{(1)}}\right]$, this

implies $E_3 = \dfrac{1}{1 - \dfrac{M_n^{(2)}}{M_n^{(1)}}} \displaystyle\int_{-\infty}^{-\frac{1}{M_n^{(1)}}} M_n^{(2)} U^{(2)}\left(M_n^{(2)}v\right) E_n(v)\,dv,$ wher

e $E_n(v) = \left(E\hat{f}_n(x-v) - f(x)\right)$

Putting $M_n^{(2)}v = u$, and using the fact that $U^{(2)}(v)$ is null outside of the interval $[-1,1]$, we obtain

$$E_3 = \frac{1}{1 - \dfrac{M_n^{(2)}}{M_n^{(1)}}} \int_{\frac{M_n^{(2)}}{M_n^{(1)}}}^{1} U^{(2)}(v)\left(E\hat{f}_n\left(x + \frac{v}{M_n^{(2)}}\right) - f(x)\right) dv.$$

Similarly, we show that

$$E_1 = \frac{1}{1 - \dfrac{M_n^{(2)}}{M_n^{(1)}}} \int_{\frac{M_n^{(2)}}{M_n^{(1)}}}^{1} U^{(2)}(v)\left[E\left(\hat{f}\left(x - \frac{v}{M_n^{(2)}}\right)\right) - f(x)\right] dv.$$

The last expressions of E_1 and E_3 can be regrouped as follows:

$$E' = \frac{1}{1 - \dfrac{M_n^{(2)}}{M_n^{(1)}}} \int_{\frac{M_n^{(2)}}{M_n^{(1)}}}^{1} U^{(2)}(v)\left[E\left(\hat{f}\left(x \pm \frac{v}{M_n^{(2)}}\right)\right) - f(x)\right] dv.$$

Showing that E' converges to 0. We show that

$$E\hat{f}\left(x \pm \frac{v}{M_n^{(2)}}\right) - f(x) = R_n' + T_n'$$

with $R_n' = \dfrac{1}{b_n}\displaystyle\int K\left(\dfrac{x \pm \dfrac{v}{M_n^{(2)}} - z}{b_n}\right) f(z)\,dz - f(x)$

and $T_n' = \dfrac{1}{b_n}\displaystyle\sum_{j=1}^{m} a_j' K\left(\dfrac{x \pm \dfrac{v}{M_n^{(2)}} - \lambda_m}{b_n}\right) f(\lambda_m).$

The term R_n', can be written as

$$R_n' = \int K(z) f\left(x \pm \frac{v}{M_n^{(2)}} - b_n z\right) dz - f(x).$$

Using Taylor's formula, we obtain:

$$f\left(x \pm \frac{v}{M_n^{(2)}} - b_n z\right) = f(x) + \left(\pm \frac{v}{M_n^{(2)}} - b_n z\right) f'(x) +$$
$$\left(\pm \frac{v}{M_n^{(2)}} - b_n z\right)^2 f''\left(x + \tau\left(\pm \frac{v}{M_n^{(2)}} - b_n z\right)\right),$$

where τ is a real number verifying $0 < \tau < 1$. K is an even density of probability. The derivatives f' and f'' are bounded and $\int z^2 K(z) < \infty$, therefore

$$R_n' = O\left(\frac{1}{M_n^{(2)}} + \left(\frac{1}{M_n^{(2)}}\right)^2 + b_n^2\right).$$

$M_n^{(2)}$ tends to infinity and $M_n^{(2)} b_n$ tends to zero, hence

$R_n' = O\left(\dfrac{1}{M_n^{(2)}}\right)$. The fact that $U^{(2)}$ is bounded, we have

$$\frac{1}{1 - \dfrac{M_n^{(2)}}{M_n^{(1)}}} \int_{\frac{M_n^{(2)}}{M_n^{(1)}}}^{1} U^{(2)}(v) R_n' = O\left(\frac{1}{M_n^{(2)}}\right)$$

For the term T_n', we distinguish two cases:

a) Let $x = \lambda_m$ with $m \in \{1,...,q\}$. Since

$\dfrac{1}{M_n^{(1)}} \le \dfrac{v}{M_n^{(2)}} \le \dfrac{1}{M_n^{(2)}}$, K is even and decreasing, we

obtain : $T_n' = O\left(\dfrac{1}{b_n} K\left(\dfrac{1}{b_n M_n^{(1)}}\right)\right).$

b) Let $x \ne \lambda_m$, for all $m \in \{1,...,q\}$, it is clair that

$$T_n' = O\left(\frac{1}{b_n} K\left(\frac{1}{b_n M_n^{(1)}}\right)\right).$$

2) Now showing that the variance converges to zero. We have

$$var(\hat{g}_n(x)) = \frac{1}{\left(1 - \dfrac{M_n^{(2)}}{M_n^{(1)}}\right)^2} E\left[\sum_{k=1}^{2} J(k) - E(J(k))\right]^2$$

with $J(k) = \displaystyle\int_{\frac{M_n^{(2)}}{M_n^{(1)}}}^{1} U^2(v) \hat{f}\left(x + (-1)^k \frac{v}{M_n^{(2)}}\right) dv.$

From Cauchy-Schwartz's inequality, we show that

$$var\left(\hat{f}\left(x + (-1)^k \frac{v}{M_n^{(2)}}\right)\right) \le \frac{1}{nb_n^2} E\left(W_{b_n}^2\left(\frac{x + (-1)^k \dfrac{v}{M_n^{(2)}} - Y_1}{b_n}\right)\right).$$

On the other hand we show that

$$\frac{1}{nb_n^2} E\left(W_{b_n}^2\left(\frac{x + (-1)^k \dfrac{v}{M_n^{(2)}} - Y_1}{b_n}\right)\right) =$$
$$\frac{1}{nb_n} \int W_{b_n}^2(z) f * h\left(x + (-1)^k \frac{v}{M_n^{(2)}} - b_n z\right) dz +$$

$$\frac{1}{nb_n}\sum_{m=1}^{p}a_m\int W_{b_n}^2(z)h\left(x+(-1)^k\frac{v}{M_n^{(2)}}-\lambda_m-b_nz\right)dz$$

$$\leq\frac{cste}{nb_n}\left\|W_{b_n}\right\|_2^2+\frac{cste}{nb_n}\sum_{m=1}^{p}a_m\left\|W_{b_n}\right\|_2^2$$

$$\leq\left(\frac{cste}{nb_n}+\frac{cste}{nb_n}\sum_{m=1}^{p}a_m\right)\left\|W_{b_n}\right\|_2^2.$$

From lemma 1, we get

$$var\left(\hat{f}\left(x+(-1)^k\frac{v}{M_n^{(2)}}\right)\right)\leq\frac{cste}{nb_n^{2\beta+1}}\left(1+\sum_{m=1}^{p}a_m\right)D(1+o(1)).$$

Thus, $var(\hat{g}(x))=O\left(\dfrac{1}{nb_n^{2\beta+1}}\left(1+\sum_{m=1}^{p}a_m\right)\right).$

III. ESTIMATION OF THE AMPLITUDES OF JUMP POINTS

Let $a(x)$ a real function defined by:

$$a(x)=\begin{cases}0 \text{ if for all } m\in\{1,...,q\}\ x\neq\lambda_m\\ a_m \text{ if } \exists m\in\{1,...,q\}\ x=\lambda_m\end{cases}$$

Choosing K such that $K(0)\neq0$ and $\lim_{y\to\infty}\left|y^2K(y)\right|=0$. Consider the following estimator

$$\hat{a}_n(x)=\frac{1}{K(0)}\sum_{i=1}^{n}\frac{1}{n}W_{h_n}\left(\frac{x-Y_j}{h_n}\right),\qquad\text{where the}$$

function W_{h_n} is defined in (3) and h_n is a real positive sequence converging to zero. The objective of the following theorem is to show that $\hat{a}_n(x)$ is an asymptotically unbiased estimator of $a(x)$.

Theorem 2 *For all $x\in R$*

1) $\lim_{n\to+\infty}E(\hat{a}_n(x))=a(x)$

2) If f is twice differentiable and its partial derivatives are continuous and bounded, then

$$E(\hat{a}_n(x))=a(x)+O(h_n).$$

Theorem 3 *Suppose that $\Phi_h(t)$ and $\Phi_K(t)$ are twice differentiable and their partial derivatives are continuous and bounded, such that*

a) $\left|\Phi_h(t)\right|>0$

b) $\lim_{t\to+\infty}t^\beta\Phi_h(t)=B$, *for some $\beta\geq1$, and $B\neq0$,*

c) $\delta_{\beta,1}\int|t|^{\beta-2}\left|\Phi_K(t)\right|dt<\infty,\ \int|t|^{\beta-1}\left|\Phi_K'(t)\right|dt<\infty,$
$\int|t|^\beta\left|\Phi_K''(t)\right|dt<\infty.$ *If $f*h$ is continuous at x and at $x-\lambda_m$ for $m=1,...,q$, then*

$$\lim_{n\to+\infty}nh_n^{2\beta-1}var(\hat{a}_n(x))=D\left(f*h(x)+\sum_{m=1}^{q}a_mh(x-\lambda_m)\right)w$$

here $D=(1/2\pi B)^2\int_{-\infty}^{+\infty}|t|^{2\beta}\left|\Phi_h(t)\right|^2dt.$

Proof: From the above results, we get

$$var(\hat{a}_n(x))=\frac{1}{nK^2(0)}E\left(W_{h_n}^2\left(\frac{x-Y_1}{h_n}\right)\right)+O\left(\frac{1}{n}\right).$$

Putting

$$A_{n1}=\frac{1}{nK^2(0)}E\left(W_{h_n}^2\left(\frac{x-Y_j}{h_n}\right)\right).$$

$$=\frac{1}{nK^2(0)}\int W_{h_n}^2\left(\frac{x-z}{h_n}\right)f*h(z)dz+$$

$$\frac{1}{nK^2(0)}\sum_{m=1}^{q}a_m\int W_{h_n}^2\left(\frac{x-\lambda_m-z}{h_n}\right)h(z)dz$$

The functions $f*h$ and h are continuous at x and at $x-\lambda_m$ for all $m\in\{1,...,q\}$, we obtain

$$\lim_{n\to+\infty}nh_n^{2\beta-1}A_{n1}=D\left(f*h(x)+\sum_{m=1}^{q}a_mh(x-\lambda_m)\right).$$

Since $h_n^{2\beta-1}$ tends to zero, we deduce

$$\lim_{n\to+\infty}nh_n^{2\beta-1}var(\hat{a}_n(x))=D\left(f*h(x)+\sum_{m=1}^{q}a_mh(x-\lambda_m)\right).$$

IV. CONCLUSION

We have presented in this paper some results about the deconvoltion estimation of the density of the continuous part and the estimation of the amplitude of jump point. This work could be applied to other cases when the distribution contains

points of discontinuity that risks being badly treated by sharing interval distribution or by using Monte Carlo method. The proposed methods can be extended to other applications in several sectors. Indeed, the control of the quality for a product manufactured in the auto industry measures the pollution of automobile. The distribution can follow a continuous law except some observations which are taken when there is fog and reached the constant value (point of the jump). In oceanography when we observe, by using a camera placed at a certain depth in water the length of the fishes. The distribution may represent some jumps due to the acceleration of movement during the passage of a predator. In Astronomy the repeated passage of an object that prevent the vision of stars (cloud, bird, ...) can create a jump of data. This work could be supplemented by the study of the central limit theorem and by giving a test for verifying the existence jump points and their localizations as well as the study of optimal smoothing parameters using cross validation techniques that have proven in this field.

References

[1]   J. Fan, On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems, 19, 3, 1109-1680 (1991).

[2]   J. Fan, Global behavior of deconvolution kernel estimates,Statistica Sinica, 1, 541-551 (1991).

[3]   R.J. Carroll and P. Hall,Optimal rates of convergence for deconvolving a density, J. Amer. Statist. Assoc., 83 : 1184-1186 (1988).

[4]   R.J. Carroll , D. Ruppert, and L. A. Stefanski, Measurement Error in Nonlinear Models, Chap-man and Hall (1995).

[5]   E. Masry, Multivariate probability density deconvolution for stationary random processes, IEEE Transactions on Information Theory, 37,4, 1105-1115 (1991).

[6]   E. Masry, Asymptotic normality for deconvolution estimators of multivariate densities of stationary processes, Journal of multivariate analysis, 44,1 : 1105-1115 (1993a).

[7]   E. Masry, Strong consistency and rates for deconvolution of multivariate densities of stationary processes. Stochastic Processes Applic., 47 : 53-74 (1993b).

[8]   P. Medgyessy,Decomposition of Superpositions of Density Functions and Discrete Distributions, placeStateNew York : Wiley (1977).

[9]   M. Rachdi and R. Sabre, Mixed-spectra analysis for stationary random fields, Statistical Methods and Applications, 18, 333-358 (2009).

[10] R. Sabre, Spectral density estimation for stationary stable random fields, Journal Applicationes Mathematicae, 23,2: 107-133 (1995).

[11] R. Sabre, Nonparametric density estimation of continuous part of a mixed measure, Statistica, 63,3 : 441-467 (2003).

[12] R. Sabre,Tests and Asymptotic Normality for certain mixture of bivariate measure, Statistica, LXX, n2, 115-136 (2010).

[13] L. Stefanski, and R.J. Carroll, deconvoluting kernel density estimators, Statistics, 21,2: 169-184 (1990).

[14] M. Van der Baan and D. Pham, Robust wavelet estimation and blind deconvolution of noisy surface seismics, Geophisics,73,5, 37–46 (2008).

[15] X. Wang, and B. Wang, Deconvolution Estimation in Measurement Error Models, Journal of Statistical Software, 39, Issue 10, 1-24 (2011).