Exponential Inequalities in Functional Nonparametrics Regression for Mixing Process

K. Belaide*
Department of Mathematics, Univ. A/Mira Bejaia, Algeria.

* Corresponding author. Email: k_tim2002@yahoo.fr
Manuscript submitted May 2, 2017; accepted September 13, 2017.
doi: 10.17706/ijapm.2017.7.4.251-258

Abstract: This paper establishes exponential inequalities for the probability of the distance between kernel estimator and its means in nonparametric regression problem with mixing variables. We consider an operator equation taking the following form \( Y = A\theta(Z) + \epsilon \), where \( A \) is a compact operator.

The goal is to estimate the functional \( \theta \) when the variable \( Z \) is contaminated by measurements errors.

Key words: Convolution linear compact operator, kernel estimator, mixing process, non parametric regression.

1. Introduction

The nonparametric estimation of function is an important tool for analyzing data, as well as inferential statistics in graphical visualizations. To this end, confidence interval and uniform confidence bands are often used in statistics. Starting with the work of Bickel and Resenblatt [1], who built confidence bands for the kernel estimator of a density function of independent and identically distributed observations. Since then, their methods were much developed in the context of estimation of density function and regression function. M. Birke, N. Bissantz and H. Holzman [2] Have constructed uniform confidence intervals for an inverse regression function in a similar way to Bickel and Rosenblatt, based on approximations and the limit theorem of a stationary Gaussian process. Recently, L. Desole [3] studies a generalization of the Nadaraya-Watson type estimators in the case of the \( \alpha \)-mixing functional variables.

In the context of inverse problems, such as calibration (inverse regression) estimator deconvolution kernel and Nadaraya-Watson are often used.

In this work, we consider a convolution operator equation of the form

\[
Y = A\theta(Z) + \epsilon
\]  

(1)

The goal is to estimate the functional \( \theta \) when the variable \( Z \) is contaminated by measurement errors.

\( A: \mathbb{H} \to \mathbb{H} \) is a convolution linear compact operator, given by

\[
A(s) = \int \Psi(s - t)\theta(t)\,dt
\]  

(2)

where \( \Psi \) is a known density function.

2. Notations and Preliminaries
We consider a sample of size \( n \), \((Y_1, Z_1), \ldots, (Y_n, Z_n)\) of variable \((Y,Z)\) satisfying the equation (1), assumed \( \alpha \)-mixing.

Assume that the sequence of regression errors \((\varepsilon_k)_{k \geq 1}\) are identically distributed random variables with zero mean and finite variance \( \sigma^2 \). We assume that for all \( k=1, \ldots, n \), \( Z_k \) is given by \( X_k = Z_k + \delta_k \) where \( \delta_k \) is a mistake to contamination.

We denote by \( \varphi_X \) the characteristic function of the variable \( X \) or the Fourier transform of the density function and \( \varphi \) is the density function of \( -\delta \).

The sequence of random variables \((Y_k)_{k \geq 1}\) is assumed to be mixing. This amounts to assume that the sequence of random variables \((\varepsilon_k)_{k \geq 1}\) is mixing.

3. Position of the Problem

Consider the model given by equation (1), where the operator \( A \) is the operator of convolution with the assumed known density function \( \psi \). The problem considered is equivalent to

\[
Y_k = \theta(X_k + \varepsilon_k) \quad \text{and} \quad X_k = Z_k + \delta_k
\]

We suppose that the random variables \( \delta_k, Z_k \) and \( \varepsilon_k \) are independent of each other. The regression error \( \varepsilon_k \) satisfies \( \mathbb{E}(\varepsilon_k/X_k) = 0 \). The regression function \( \theta \) is such that \( \theta(x) = \mathbb{E}(Y_k/X_k = x) \) Related to the observed regression function \( g(z) = \mathbb{E}(Y_k/Z_k = z) \) by \( g(z) = \theta \ast \psi(z) \). Where \( \psi \) is the density function of \( -\delta \).

In practice, two essential cases arise. The observations \( Z_k \) are either fixed points (ex. M. Birke, N. Bissantz et H. Holzman [2]) or \( Z_k \) are random variables (ex. L. Desol [4], A. Tadj [5]). In this work, we will consider the case where \( Z_k = z_k \) are fixed point.

The model considered is a classical nonparametric regression model with deconvolution (see [6]).

We propose a kernel type estimator for the functional \( \theta \), we assume that the Fourier transform of the density function \( \varphi \) is such that \( \varphi(\omega) \neq 0 \) for all \( \omega \in \mathbb{R} \) and the Fourier transform of the kernel \( k \) has compact support. The deconvolution kernel estimator of \( \theta \) is given by

\[
\hat{\theta}_n(x) = \frac{1}{\pi} \int \exp(-i \omega x) \frac{\varphi_k(\omega) \varphi(\omega)}{\varphi(\omega)} \, d\omega
\]

where \( h>0 \) is a smoothing parameter and \( \varphi_k(\cdot) \) is the empirical Fourier transform of \( g \) given by

\[
\varphi_k(\omega) = \frac{1}{n \alpha_n} \sum_{k=1}^{n} Y_k \exp(i \omega Z_k)
\]

We have \( g = \theta \ast \psi \) and its estimator

\[
g_n(x) = \hat{\theta}_n \ast \Psi(x)
\]

It’s clear that

\[
g_n(x) = \frac{1}{n \alpha_n h} \sum_{k=1}^{n} Y_k k \left( \frac{x-z_k}{h} \right)
\]

In this case, one can easily see that the estimator \( \hat{\theta}_n(x) \) of \( \theta(x) \) given (3) is written in the kernel form as
\[ \theta_n(x) = \frac{1}{nha_n} \sum_{k=1}^{n} Y_k K(\frac{x - z_k}{h}) \]

where the kernel \( K \) is given by

\[ K(x, h) = \frac{1}{2\pi} \int \exp(-i\omega x) \frac{\varphi_k(\frac{\omega}{h})}{\varphi(\frac{\omega}{h})} d\omega \]

(4)

From the estimate of the deconvolution density function, it is well known that the optimal rate for what is estimated \( \theta \) depends on \( \theta \) smoothing and that of \( \psi \) or equivalently, on the Fourier transform properties.

\( \psi \) is said to be moderately smooth and the problem is moderately ill-posed if the Fourier transform \( |\varphi(\omega)| \) decreases polynomialally when \( t \) tends to infinity. In this case, the optimal rate for estimating \( \theta \) is also polynomial. If, on the contrary, \( |\varphi(\omega)| \) decreases exponentially when \( t \) tends to infinity, \( \psi \) is said to be smoothed and the problem is said to be very badly posed. The optimal rate of convergence is logarithmic order. For more details on the estimation of density functions see Fan [7] and Pensky and Vidakovic [8].

In the following, we will restrict ourselves to an ordinary smoothing of \( \psi \) which means that the model (1) is a moderately ill-posed problem.

Assumes that

\[ \varphi(\omega) \omega^\beta \to C_\epsilon \quad \text{when} \quad \omega \to \infty \]

(5)

For some \( \beta \geq 0 \) and \( C_\epsilon \in \mathbb{C} \setminus \{0\} \).

Note that this implies that

\[ \varphi(\omega) |\omega|^\beta \to \widetilde{C}_\epsilon \quad \text{when} \quad \omega \to \infty \]

Under the hypothesis (5) and by the dominated convergence theorem [6], We have an asymptotic form of the deconvolution kernel (4) quite simple.

\[ h^\beta K(x, h) = \frac{h^\beta}{2\pi} \int_0^{+\infty} \exp(-i\omega x) \frac{\varphi_k(\frac{\omega}{h})}{\varphi(\frac{\omega}{h})} d\omega + \frac{h^\beta}{2\pi} \int_{-\infty}^0 \exp(-i\omega x) \frac{\varphi_k(\omega)}{\varphi(\omega)} d\omega \]

\[ = L(x) \]

It’s clear that \( L(x) \in \mathbb{R} \). This result will be used in the calculation of the variance of \( \theta_n(x) \).

**Hypothesis**

H1: The Fourier transform \( \varphi_k \) of \( k \) est symétrique.

H2: \( \int K(z,h)|z|^{(3/2)}(\log\log^*|z|)^{(1/2)}dz=O(h^\beta) \) with \( \log\log^*|z|=0 \) if \( |z|<e \) and \( \log\log^*|z|=\log\log|z| \) somewhere else.

H3: \( \theta \) is j-time differentiable and \( h^j = O \left( \frac{1}{\sqrt{n|h|^{1/2}}} \right) \) equivalent to \( h = O \left( n^{-\frac{1}{2j+2}} \right) \)

H4: \( (Y_i)_{i\geq 1} \) is \( \alpha \)-mixing.

H5: \( \exists M >0 \), suth that \( |\epsilon K(x,h)|<M \).

H6: \( \max \{ E(|Y_k Y_l|/X_k X_l), E(|Y_k|/X_k X_l) \} \leq C \) a.s.
H7 : \( \exists (u_n) \in \mathbb{N}^\infty, o(u_n) + o(n[\alpha(u_n)])^{(p-2)/2} \to 0 \).

The following lemma gives a simple asymptotic expression for the variance of \( \theta_n(x) \).
The following lemma gives the asymptotic expression of the variance of \( \theta_n(x) \), in the case of strong mixing variables.

**Lemma** : Under the assumptions H4 to H7, we have,

\[
\sum_{k=1}^{n} \sum_{l=1, l \neq k}^{n} \text{cov}(Y_k, Y_l) K\left(\frac{x - z_k}{h}, h\right) K\left(\frac{x - z_l}{h}, h\right) = o\left(\sum \text{Var}(Y_k) K\left(\frac{x - z_k}{h}, h\right)\right)
\]

**Proof**

\[
\sum_{k=1}^{n} \sum_{l=1, l \neq k}^{n} \text{cov}(Y_k, Y_l) K\left(\frac{x - z_k}{h}, h\right) K\left(\frac{x - z_l}{h}, h\right) = \sum_{1<|k-l|<u} \sum_{u_n<|k-l|<n} \left| \text{cov}(\xi_1, \xi_2) \right| \leq c |\xi_1|_p |\xi_2|_q
\]

where \( c \) is a positive constant.

Consider \( p = q = 2 \); This implies that \( r = \frac{p}{p-2} \). It comes

\[
|B| \leq c |\xi_1|_p |\xi_2|_q \sum_{u_n<|k-l|<n} |\alpha(k-l)|^{p-2} \frac{1}{2\pi} \int \left| \exp\left(-i\omega \frac{x - z_k}{h}\right) \right| \frac{\phi_k(\omega)}{\Phi(\omega)} d\omega
\]

\[
h^2|B| \leq c |\xi_1|_p |\xi_2|_q \sum_{u_n<|k-l|<n} |\alpha(k-l)|^{p-2} \frac{1}{2\pi} \left( \int \frac{\phi_k(\omega)}{\Phi(\omega)} d\omega \right)^2
\]

Using hypothesis (5) and positing \( C_0 = c |\xi_1|_p |\xi_2|_q \)

We have

\[
|B| \leq C_0 \sum_{u_n<|k-l|<n} |\alpha(k-l)|^{p-2} \frac{1}{2\pi} \int_0^{+\infty} \omega^{\alpha} \frac{\phi_k(\omega)}{\Phi(\omega)} d\omega + \frac{1}{2\pi} \int_{-\infty}^{0} |\omega|^{\beta} \frac{\phi_k(\omega)}{\Phi(\omega)} d\omega
\]
\[ \rightarrow C_0 k^2(0)[\alpha(k - l)]^\frac{p-2}{p}(n - u_n - 1)(n - u_n) \]

Let us now return to the term \( A \)

\[ |A| \leq \sum_{1 < |k - l| < u_n} |E(E(Y_k Y_l / Y_k) )K \left( \frac{x - z_k}{h}, h \right) K \left( \frac{x - z_l}{h}, h \right)| \]

\[ \leq c \sum_{1 < |k - l| < u_n} \left| K \left( \frac{x - z_k}{h}, h \right) K \left( \frac{x - z_l}{h}, h \right) \right| \]

\[ \leq n u_n c \max_{1 < |k - l| < u_n} \left( \left[ \frac{1}{2\pi} \int_{\omega} \frac{\phi_h(\omega)}{\phi_f(\omega)} d\omega \right]^2 \right) \]

By hypothesis (5),

\[ h^{2\beta} |A| \leq n u_n c \max_{u_n c |k - l| < n} \left( \left[ \frac{1}{2\pi} \int_{\omega} \frac{\phi_h(\omega)}{\phi_f(\omega)} d\omega \right]^2 + \int_{-\infty}^{0} \frac{\phi_h(\omega)}{\phi_f(\omega)} |\omega|^{\beta} d\omega \right) \]

When \( h \) tends to 0, we have

\[ h^{2\beta} \sum_{k=1}^{n} \sum_{l=1, l \neq k}^{n} \text{cov}(Y_k, Y_l) K \left( \frac{x - z_k}{h}, h \right) K \left( \frac{x - z_l}{h}, h \right) \]

\[ \leq n u_n c K^2(0) + C_0 k^2(0)[\alpha(k - l)]^\frac{p-2}{p}(n - u_n - 1)(n - u_n) \]

\[ \leq n C_1 K^2(0)(u_n + n)[\alpha(k - l)]^\frac{p-2}{p} \]

On the other hand

\[ \sum_{k=1}^{n} \text{Var} \left( Y_k K \left( \frac{x - z_k}{h}, h \right) \right) = \sum_{k=1}^{n} \text{Var}(Y_k) K^2 \left( \frac{x - z_k}{h}, h \right) \]

\[ = \sum_{k=1}^{n} \text{Var}(\varepsilon_k) K^2 \left( \frac{x - z_k}{h}, h \right) = \frac{\sigma^2}{h^{2\beta}} \sum_{k=1}^{n} h^{2\beta} K^2 \left( \frac{x - z_k}{h}, h \right) \]

For \( h \to 0 \), \( \sum_{k=1}^{n} \text{Var} \left( Y_k K \left( \frac{x - z_k}{h}, h \right) \right) \) is asymptotically equal to \( n \sigma^2 K^2(0) \)

And

\[ \sum_{k=1}^{n} \sum_{l=1, l \neq k} \text{cov}(Y_k, Y_l) K \left( \frac{x - z_k}{h}, h \right) K \left( \frac{x - z_l}{h}, h \right) \]

\[ \leq \frac{C_1}{\sigma^2} (u_n + n)[\alpha(k - l)]^\frac{p-2}{p} \to 0, \text{when } n \to \infty \]

This shows that the variance of \( \theta_n(x) \) is asymptotically proportional to
Then
\[
\text{Var}(\theta_n(x)) = \frac{1}{(n/\alpha_n)^2} \sum_{k=1}^{n} \text{Var} \left( Y_k K \left( \frac{x - z_k}{h} , h \right) \right)
\]
\[+ \frac{1}{(n/\alpha_n)^2} \sum_{k=1}^{n} \sum_{l \neq k} \text{cov} \left( Y_k K \left( \frac{x - z_k}{h} , h \right), Y_l K \left( \frac{x - z_l}{h} , h \right) \right) = \frac{1}{(n/\alpha_n)^2} S_1 + \frac{1}{(n/\alpha_n)^2} S_2
\]

With $S_1$ is asymptotically proportional to $n\sigma^2$
And $S_2 = o \left( \sum_{k=1}^{n} \text{Var} \left( Y_k K \left( \frac{x - z_k}{h} , h \right) \right) \right)$
Thus, we obtain an asymptotic expression of the variance of the estimator $\theta_n(x)$
\[
\text{Var}(\theta_n(x)) = \frac{\sigma^2}{n^2 h^2 + 2\beta} a_n^2 K^2(0) + \frac{1}{n^2 h^2 a_n^2} o \left( \frac{n\sigma^2}{h^2} K^2(0) \right)
\]

4. Exponential Inequality
We will establish an exponential inequality of the probability of deviation of the estimator $\theta_n(x)$ given by (3) of $\theta(x)$ to its mathematical expectation.
We put
\[
Z_n(x) = \frac{\sqrt{n h + \beta a_n}}{\sigma} [\theta_n(x) - E(\theta_n(x))]
\]

Theorem: Under some assumptions, for all $\epsilon > 0, \forall (n \geq 4) \forall k \in \{1, \ldots, [(n/2) - 1] \} \forall \epsilon \in [0, \ldots, (4kMe)^{-1}]$, we have
\[
P(\{|Z_n(x)| > \epsilon\}) \leq 2\exp \left( -n \eta \left( \frac{\sigma \epsilon}{h^2 \sqrt{n}} - 6 \eta e(V + 8M^2\Sigma a_i) - \left( 2\sqrt{\frac{\epsilon}{k \eta}} \right)^2 e \right) \right)
\]

Proof
\[
Z_n(x) = \frac{\sqrt{n h + \beta a_n}}{\sigma} \left[ \frac{1}{n h a_n} \sum_{k=1}^{n} Y_k K \left( \frac{x - z_k}{h} , h \right) - \frac{1}{n h a_n} \sum_{k=1}^{n} E(Y_k) K \left( \frac{x - z_k}{h} , h \right) \right]
\]
\[
Z_n(x) = \frac{h \beta}{\sqrt{n} \sigma} \left[ \sum_{k=1}^{n} (Y_k - E(Y_k)) K \left( \frac{x - z_k}{h} , h \right) \right]
\]
\[
Z_n(x) = \frac{h \beta}{\sqrt{n} \sigma} \left[ \sum_{k=1}^{n} \xi_k K \left( \frac{x - z_k}{h} , h \right) \right]
\]

From the exponential type inequality for the mixed random variables, the Carbon inequality [9] applied to centred $\alpha$ – mixing variables for $\xi_k = \xi_k K \left( \frac{x - z_k}{h} , h \right)$
We have $E(\xi_k) = 0$
\[ P(\left| \xi_k \right| > \varepsilon) \leq 2\exp\left( -n \frac{\eta}{\sqrt{n}} \left( \frac{\sigma_x}{h^\beta \sqrt{n}} - 6 \eta e (V + 8M^2 \sigma_i) - \left( \frac{2\sqrt{\varepsilon}}{k \eta} \right)^{\frac{2\varepsilon}{3n}} \right) \right) \]

\[ P\left( |Z_n(\alpha)| > \frac{h^\beta}{\sqrt{n}} \varepsilon \right) \leq 2\exp\left( -n \frac{\eta}{\sqrt{n}} \left( \frac{\sigma_x}{h^\beta \sqrt{n}} - 6 \eta e (V + 8M^2 \sigma_i) - \left( \frac{2\sqrt{\varepsilon}}{k \eta} \right)^{\frac{2\varepsilon}{3n}} \right) \right) \]

We put, \( \varepsilon' = \frac{h^\beta}{\sqrt{n} \sigma_x} \varepsilon \), we obtain

\[ P(\left| Z_n(\alpha) \right| > \varepsilon') \leq 2\exp\left( -n \frac{\eta}{\sqrt{n}} \left( \frac{\sigma_x}{h^\beta \sqrt{n}} - 6 \eta e (V + 8M^2 \sigma_i) - \left( \frac{2\sqrt{\varepsilon}}{k \eta} \right)^{\frac{2\varepsilon}{3n}} \right) \right) \]

Corollary: Under assumptions of the theorem, and if in addition, the mixing coefficients \( \alpha_k \) satisfy \( \alpha_k \leq a^k \) for \( a>0 \) and \( 0<\rho<1 \), then, for all \( \gamma \in ]0,1[ \), there exists \( \eta_\gamma \) such that for all \( n \geq \eta_\beta \), we have

\[ P\left( \frac{1}{n} |Z_n(\alpha)| > \varepsilon \right) \leq 2\exp\left( -\frac{n^{1-\gamma/2}}{h^5 \sigma_x} \right) \]

See the corollary of the inequality of large deviation given by Carbone [10]

5. Conclusion

Under the assumptions of theorem, the sequence \( (\theta_n) \) converges almost completely (a.co.) to the exact solution \( \theta \) of the equation (1).

Under the assumptions of Theorem, we have

\[ \theta_n - \theta = o\left( \frac{\log(n)}{\sqrt{n}} \right) \]

Under the assumptions of Theorem, for a given significance threshold \( \gamma \), it exists an integer \( n_\gamma \) for which

\[ P(|\theta_n - \theta| \leq \varepsilon) > 1 - \gamma \]

ie, the exact solution \( \theta \) of equation (1) belongs to the closed ball with center \( \theta_n \) and radius \( \varepsilon \) with probability greater than or equal to \( 1-\gamma \).

Exponential inequality helps build confidence bands of the deconvolution kernel estimator \( \theta_n(x) \).

References


K. Belaide was born in Bejaia, Algeria. After obtained magister from Sciences and technology Houari Boumedienne University in Algiers Algeria, she was recruited as a master assistant at A / Mira University in Bejaia and enrolled in PhD in Mathematic option Analysis and Probabily. A dissertation thesis of her thesis was in 2013 and Habilitation in 2015. She works actually in Laboratory of Applied Mathematics (LMA), in Stochastic inverse problems, and fractionary processes. She teaches in mathematics since 1995 to this day and Supervision of students Ingenieur, Licences, Master and PhD. She published 2 papers in refereed international journals. She has been involved in many international conferences as tutorial presenter.