# Exponential convergence rates of kernel density functio for negatively associated random variables

Jabbari Nooghabi, H.
Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran.
E-mail: Jabbarinh@um.ac.ir

Abstract

Under mild conditions on the covariance structure of the sample, we estimate densit function of negatively associated random variables based on kernel estimator and prov exponential rates for this estimator with a uniform version, over compact sets. The proo uses a block decomposition of the sums to allow an approximation to independence. Som examples supposing exponential but also polynomial rates on the covariances that fulfil our assumptions are presented in the last section.

*AMS Subject Classification:* 62G07; 62G20

*Keywords:* Covariance function; Exponential Rates; Kernel estimation; Negativ Association.

## 1. Introduction

There exist several versions available in the literature for independent sequences c variables with assumptions of uniform boundedness or some, quite relaxed, control on thei (centered or noncentered) moments. This independence assumption was eventuall replaced by some kind of control on the dependence structure of the sample upon which th estimation is carried. There was various estimation methods proposed, among which w will be interested in the nonparametric kernel estimator. For an account of results an literature we refer the reader to Bosq [5].

One of the dependence structures that has attracted the interest of probabilists an statisticians is negative association. Some definitions and applications as introduced b Alam and Saxena [1] and carefully studied by Joag-Dev and Proschan [8] and Block et a [4]. The significance of exponential inequalities toward several probability and statistica applications is well known. The consistency of the kernel estimator under associate sampling was proved by Bagai and Prakasa Rao [3] which returned to the problem in Baga and Prakasa Rao [2] proving a uniform consistency result. Independently, Roussas [10] als proved a consistency result for the kernel estimator under associated sampling. Henrique and Oliveira [7] obtained exponential convergence rates for the kernel estimator of th density function under associated sampling. To the best of our knowledge, asymptoti behavior of kernel density function and exponential rates for its estimation under N random variables, are not available in the literature.

The remaining sections of the article are organized as follows: in the next section, definitions and the necessary notation and terminology are introduced before the preliminary result. In addition to the basic assumption of negative association, the condition required is that the underlying random variables are strictly stationary with bounded and continuous density function. In Section 3, we introduce some preliminary results. Actually, all preliminary results needed are taken care of in the same section. The proof of the main theorems rests on Lemma 4.1 and 4.2, which are formulated and proved in Section 4. In the final section we will present some examples of covariance structures that fulfill the assumptions used in this article. In this section, we will provide an example of polynomial increase rate on the covariance structure that still verifies the conditions under which the general results hold. Another example is based on geometrically rate on the covariance structure.

## 2. Definitions, assumptions and some lemmas

Let $\{X_n, n \geq 1\}$ be a sequence of random variables with the same common unknown density function $f$.

**Definition 2.1:** Let $K$ be a fixed probability density and $h_n$ a sequence of nonnegative real numbers converging to zero. Then, the kernel estimator of the density function $f$ is defined as

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n}\right),$$

which is well known to be asymptotically unbiased, if there exists a bounded and continuous version of the density. Under these assumptions on $f$, the convergence of $E[\hat{f}_n(x)]$ to $f(x)$ is uniform on compact sets.

**Definition 2.2:** Two random variables $X$ and $Y$ are negatively quadrant dependent (NQD) if for every $x, y \in \Re$ we have

$$P_r(X \leq x, Y \leq y) \leq P_r(X \leq x).P_r(Y \leq y).$$

**Definition 2.3:** A finite family of random variables $\{X_i, 1 \leq i \leq n\}$ is said to be negatively associated (NA) if for every pair of disjoint subsets A and B of $\{1,2,\ldots,n\}$,

$$Cov(g_1(X_i, i \in A), g_2(X_j, j \in B)) \leq 0,$$

whenever $g_1$ and $g_2$ are coordinatewise increasing and such that the covariance exists. An infinite family of random variables is NA if every finite subfamily is NA.

A technical problem arises when dealing with $\hat{f}_n(x)$ for negatively associated variables. In fact, association is only preserved under monotone transformations, which means that, in general, the variables $K\left(\frac{x - X_1}{h_n}\right), K\left(\frac{x - X_2}{h_n}\right), \ldots$ are not associated. This problem is resolved, as usual, by supposing the kernel $K$ to be bounded variation.

Now, we introduce a set of assumptions that need to prove the main results.

**A1.** $\{X_n, n \geq 1\}$ is strictly stationary and negatively associated random variables with common continuous density function $f$ which is bounded by a positive constant $M_0$, that is $M_0 = \sup_{x\in\Re} |f(x)|$.

**A2.** The kernel function $K$ is a probability density of bounded variation such that $\int K^2(u)\,du < \infty$; further, if $K = K_1 - K_2$ where $K_1$ and $K_2$ are nondecreasing functions, the derivatives $K_1'$ and $K_2'$ exist and are integrable.

**Remark 1.** Under **A1**, we have that

$$\left|F_{X_1,X_j}(r,s) - F_{X_1}(r).F_{X_j}(s)\right| \leq M_1 \left|Cov^{1/3}(X_1, X_j)\right|, \quad r,s \in \Re, \tag{2.1}$$

where $F_{X_1,X_j}$ and $F_{X_1}$ represent the distribution functions of $(X_1, X_j)$ and $X_1$, respectively, and $M_1 = 2Max(2/\pi^2, 45M_0)$ (see corollary of Theorem 1 in Sadikova [11] and relation (21) in Newman [9] for details). This inequality provides an upper bound for the covariances between the variables $K_q\left(\frac{x - X_j}{h_n}\right)$, $q = 1,2,\ j = 1,2,\ldots$.

**Lemma 2.1.** Suppose $U$ and $V$ are NQD random variables with finite variance and $g_1, g_2$ are complex valued functions on $\Re^1$ with $g_1'$ and $g_2'$ bounded. Then

$$|Cov(g_1(U), g_2(V))| \leq -\|g_1'\|_\infty . \|g_2'\|_\infty \, Cov(U,V),$$

where $\|.\|_\infty$ denotes the sup norm on $\Re^1$; in particular, for any real $r, s$,

$$\left|E\left(e^{irU+isV}\right) - E\left(e^{irU}\right)E\left(e^{isV}\right)\right| \leq -|r|.|s| Cov(U,V).$$

**Proof.** Define

$$H(r,s) = P_r(U \leq u, V \leq v) - P_r(U \leq u).P_r(V \leq v).$$

By Hoeffding lemma,

$$Cov(U,V) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(u,v)\,dudv.$$

This equation can be easily generalized to yield

$$Cov(g_1(U), g_2(V)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1'(u).g_2'(v) H(u,v)\,dudv;$$

Since $U$ and $V$ are NQD random variables $H(u,v) \leq 0$, thus

$$|Cov(g_1(U), g_2(V))| \leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |g_1'(u)|.|g_2'(v)|.|H(u,v)|\,dudv$$
$$\leq -\|g_1'\|_\infty . \|g_2'\|_\infty \, Cov(U,V),$$

as desired. □

**Lemma 2.2.** Suppose $\{X_n, n \geq 1\}$ satisfy **A1** and the kernel function satisfy **A2**. Then,

$$\left|Cov\left(K_q\left(\frac{x - X_1}{h_n}\right), K_q\left(\frac{x - X_j}{h_n}\right)\right)\right| \leq M_1 \left(\int K_q'(r)\,dr\right)^2 \left|Cov^{1/3}(X_1, X_j)\right|, \quad q = 1,2.$$

**Proof.** Just notice that

$$\left|Cov(K_q\left(\frac{x-X_1}{h_n}\right), K_q\left(\frac{x-X_j}{h_n}\right))\right| =$$
$$\frac{1}{h_n^2}\int\left|K_q'\left(\frac{x-r}{h_n}\right)K_q'\left(\frac{x-s}{h_n}\right)(F_{X_1,X_j}(r,s)-F_{X_1}(r).F_{X_j}(s))\right|drds$$

and apply (2.1). □

**Lemma 2.3.** Let $X_1,\dots,X_n$ be NA random variables bounded by a constant $B$. Then, for every $\lambda>0$,

$$\left|\mathrm{E}(e^{\lambda\sum_{i=1}^{n}X_i})-\prod_{i=1}^{n}\mathrm{E}(e^{\lambda X_i})\right| \le -\lambda^2 e^{n\lambda B}\sum_{1\le i<j\le n}Cov(X_i,X_j).$$

**Proof.** By Lemma 2.1 for $\lambda>0$ we have

$$\left|Cov(e^{\lambda X_1},e^{\lambda X_2})\right| \le \lambda^2 e^{2\lambda B}\left|Cov(X,Y)\right|.$$

The results follow by induction and using the fact that if $X$, $Y$ and $Z$ are NA then so are $X$ and $Y+Z$ as they are increasing functions of NA random variables. □

We quote next a general lemma used to control some of the terms appearing in the course of proof.

**Lemma 2.4.** (Devroye, [6]). Let X be a central random variable. If there exist $a,b\in\Re$ such that $P_r(a\le X\le b)=1$, then, for every $\lambda>0$,

$$\mathrm{E}(e^{\lambda X})\le \exp(\frac{\lambda^2(b-a)^2}{8}).$$

For the formulation of the assumptions to be made in this paper, the introduction of some notation is required. Given **A2**, define

$$\widehat{f}_{1,n}(x)=\frac{1}{nh_n}\sum_{j=1}^{n}K_1\left(\frac{x-X_j}{h_n}\right), \qquad \widehat{f}_{2,n}(x)=\frac{1}{nh_n}\sum_{j=1}^{n}K_2\left(\frac{x-X_j}{h_n}\right),$$

so that $\widehat{f}_n(x)=\widehat{f}_{1,n}(x)-\widehat{f}_{2,n}(x)$. For each $n\in\mathrm{N}$, $j=1,\dots,n$ and $q=1,2$, let

$$T_{q,j,n}=h_n^{-1}(K_q\left(\frac{x-X_j}{h_n}\right)-EK_q\left(\frac{x-X_j}{h_n}\right)). \tag{2.2}$$

Note that, all these variables are monotone transformations of the initial variables $X_n$ then, they are NA.

Consider a sequence of natural numbers $p_n$ such that, for each $n\ge1$, $p_n<n/2$ and let $r_n$ be the greatest integer less or equal to $n/2\,p_n$. Define then, for $q=1,2$ and $j=1,\dots,2r_n$

$$Y_{q,j,n}=\sum_{k=(j-1)p_n+1}^{jp_n}T_{q,k,n}. \tag{2.3}$$

Note that, if the kernel $K$ satisfies **A2**, the functions $K_1$ and $K_2$ may be chosen bounded so that each variable $Y_{q,j,n}$ is bounded by $2p_n h_n^{-1}\|K_q\|_\infty$, where $\|\|$ represents the supremum norm.

Finally, for each $q=1,2$ and $n \geq 1$, define

$$Z_{q,n,od} = \sum_{j=1}^{r_n} Y_{q,2j-1,n}, \qquad Z_{q,n,ev} = \sum_{j=1}^{r_n} Y_{q,2j,n}. \tag{2.4}$$

With these definitions, if $n = 2r_n p_n$ we have $\hat{f}_{q,n}(x) - E[\hat{f}_{q,n}(x)] = \frac{1}{n}(Z_{q,n,od} + Z_{q,n,ev})$.

### 3. Some preliminary results

In this section we prove two elementary lemmas that are provided the way to the proof of the main result.

It is obvious that $|Y_{q,j,n}| \leq 2p_n h_n^{-1}\|K_q\|_\infty$, for $j=1,\dots,r_n$ and $q=1,2$. Then, we can use Lemma 2.4 to control the Laplace transform of these variables. A simple application of this lemma produces the following upper bounds.

**Lemma 3.1.** Let $X_1, X_2, \dots$ be random variables and suppose that **A2** is satisfied. If $Y_{q,j,n}$, $q=1,2$, $j=1,\dots,2r_n$ are defined by (2.3) then, for every $\lambda > 0$,

$$\prod_{j=1}^{r_n} E(e^{\frac{\lambda}{n}Y_{q,2j-1,n}}) \leq \exp\left(\frac{\lambda^2 p_n \|K_q\|_\infty^2}{nh_n^2}\right), \quad q=1,2$$

$$\prod_{j=1}^{r_n} E(e^{\frac{\lambda}{n}Y_{q,2j,n}}) \leq \exp\left(\frac{\lambda^2 p_n \|K_q\|_\infty^2}{nh_n^2}\right). \quad q=1,2$$

**Lemma 3.2.** Suppose **A1** and **A2** are satisfied. On account of (2.2), (2.3) and (2.4), and for every $\lambda > 0$,

$$\left|\mathrm{E}(e^{\frac{\lambda}{n}Z_{q,n,od}}) - \prod_{j=1}^{r_n}\mathrm{E}(e^{\frac{\lambda}{n}Y_{q,2j-1,n}})\right| \leq -\frac{\lambda^2}{2n}\exp\left(\frac{\lambda\|K_q\|_\infty}{h_n}\right)\sum_{j=p_n+2}^{(2r_n-1)p_n} Cov(T_{q,1,n}, T_{q,j,n}), \quad q=1,2 \tag{3.1}$$

and analogously for the term corresponding to $Z_{q,n,ev}$.

**Proof.** The variables defined in (2.2) are negatively associated. According to (2.3), we have, from direct application of Lemma 2.3

$$\left|\mathrm{E}(e^{\frac{\lambda}{n}Z_{q,n,od}}) - \prod_{j=1}^{r_n}\mathrm{E}(e^{\frac{\lambda}{n}Y_{q,2j-1,n}})\right| \leq -\frac{\lambda^2}{n^2}\exp\left(\frac{\lambda}{n}r_n\frac{2p_n\|K_q\|_\infty}{h_n}\right)\sum_{1\leq j<j'\leq r_n} Cov(Y_{q,2j-1,n}, Y_{q,2j'-1,n})$$

$$\leq -\frac{\lambda^2}{n^2}\exp\left(\frac{\lambda\|K_q\|_\infty}{h_n}\right)\sum_{1\leq j<j'\leq r_n} Cov(Y_{q,2j-1,n}, Y_{q,2j'-1,n}). \text{ (by } 2r_np_n \leq n) \tag{3.2}$$

Using the stationary of the variables it follows that:

$$\sum_{1\le j<j'\le r_n} Cov(Y_{q,2j-1,n}, Y_{q,2j'-1,n}) = \sum_{j=1}^{r_n-1}(r_n - j)Cov(Y_{q,1,n}, Y_{q,2j-1,n}). \tag{3.3}$$

And,

$$\begin{aligned} Cov(Y_{q,1,n}, Y_{q,2j-1,n}) &= Cov(Y_{q,1,n}, \sum_{k=(2j-2)p_n+1}^{(2j-1)p_n} T_{q,k,n}) \\ &= p_n \sum_{k=(2j-2)p_n+1}^{(2j-1)p_n} Cov(T_{q,1,n}, T_{q,k,n}) \\ &\ge p_n \sum_{k=(2j-1)p_n+2}^{(2j+1)p_n} Cov(T_{q,1,n}, T_{q,k,n}). \end{aligned} \tag{3.4}$$

Then, by inserting this into (3.3) and remind again that $r_n p_n \le n/2$, we find

$$\begin{aligned} \sum_{1\le j<j'\le r_n} Cov(Y_{q,2j-1,n}, Y_{q,2j'-1,n}) &\ge r_n p_n \sum_{j=1}^{r_n-1} \sum_{k=(2j-1)p_n+2}^{(2j+1)p_n} Cov(T_{q,1,n}, T_{q,k,n}) \\ &= r_n p_n \sum_{k=p_n+2}^{(2r_n-1)p_n} Cov(T_{q,1,n}, X_{q,k,n}) \\ &\ge \frac{n}{2} \sum_{k=p_n+2}^{(2r_n-1)p_n} Cov(T_{q,1,n}, T_{q,k,n}). \end{aligned}$$

Inserting this into (3.2), lemma follows. □

## 4. Main results

Now, we prove in position an exponential inequality for the sum of odd indexed or even indexed terms and then for the estimator. For this goal, we assumed that $r_n$ and $p_n$ are sequences such that $\frac{n}{2r_n p_n} \longrightarrow 1$. We will need to choose these sequences conveniently to prove our results.

**Lemma 4.1.** Suppose **A1** and **A2** are satisfied. Further assume that

$$-\frac{nh_n^4}{p_n^2}\exp(\frac{nh_n}{p_n}) \sum_{j=p_n+2}^{\infty} Cov(T_{q,1,n}, T_{q,j,n}) \le C_0 < \infty. \tag{4.1}$$

Then, for every $\varepsilon \in (0, \min(\|K_1\|_\infty, \|K_1\|_\infty^2, \|K_2\|_\infty, \|K_2\|_\infty^2))$ and $q = 1,2$

$$P_r(\frac{1}{n}|Z_{q,n,od}| > \varepsilon) \le (1 + C_0)\exp(-\frac{n\varepsilon^2 h_n^2}{4p_n\|K_q\|_\infty}), \tag{4.2}$$

and similarly for $Z_{q,n,ev}$.

**Proof.** Using Markov's inequality and Lemma 3.1 we find that, for every $\lambda > 0$,

$$P_r(\frac{1}{n}|Z_{q,n,od}|>\varepsilon)\le E\left(e^{\frac{\lambda}{n}Z_{q,n,od}}\right)e^{-\lambda\varepsilon}\le\left(\left|\mathrm{E}(e^{\frac{\lambda}{n}Z_{q,n,od}})-\prod_{j=1}^{r_n}\mathrm{E}(e^{\frac{\lambda}{n}Y_{q,2j-1,n}})\right|+\left|\prod_{j=1}^{r_n}\mathrm{E}(e^{\frac{\lambda}{n}Y_{q,2j-1,n}})\right|\right)e^{-\lambda\varepsilon}$$

$$\le e^{-\lambda\varepsilon}\left|\mathrm{E}(e^{\frac{\lambda}{n}Z_{q,n,od}})-\prod_{j=1}^{r_n}\mathrm{E}(e^{\frac{\lambda}{n}Y_{q,2j-1,n}})\right|+\exp\left(\frac{\lambda^2 p_n\|K_q\|_\infty^2}{nh_n^2}-\lambda\varepsilon\right) \quad (4.3)$$

Optimum value of $\lambda$ in the exponent of the last term of (4.3) is $\frac{\varepsilon nh_n^2}{2p_n\|K_q\|_\infty^2}$, so that this exponent becomes equal to $-\frac{\varepsilon^2 nh_n^2}{4p_n\|K_q\|_\infty^2}$. From Lemma 3.2 it follows, using (4.1), that

$$\left|\mathrm{E}(e^{\frac{\lambda}{n}Z_{q,n,od}})-\prod_{j=1}^{r_n}\mathrm{E}(e^{\frac{\lambda}{n}Y_{q,2j-1,n}})\right|\le C_0.$$

Then, by replacing optimum value of $\lambda$ into the first term of the upper bound of (4.3), we have

$$P_r(\frac{1}{n}|Z_{q,n,od}|>\varepsilon)\le C_0\exp\left(-\frac{\varepsilon^2 nh_n^2}{2p_n\|K_q\|_\infty^2}\right)+\exp\left(-\frac{\varepsilon^2 nh_n^2}{4p_n\|K_q\|_\infty^2}\right)$$

$$\le(1+C_0)\exp\left(-\frac{\varepsilon^2 nh_n^2}{4p_n\|K_q\|_\infty^2}\right).$$

□

In order to state the main result we have to deal with the terms in $\hat{f}_{q,n}(x)$ that are not in $\frac{1}{n}(Z_{q,n,od}+Z_{q,n,ev})$. But these are, as expected, negligible. For easier reference, define

$$R_{q,n}=\hat{f}_{q,n}(x)-E[\hat{f}_{q,n}(x)]-\frac{1}{n}(Z_{q,n,od}+Z_{q,n,ev}).\quad q=1,2 \quad (4.4)$$

**Lemma 4.2.** Let $R_{q,n}$ be defined by (4.4) for $q=1,2$. Suppose that **A1** and **A2** are satisfied and

$$\frac{nh_n^2}{p_n}\longrightarrow+\infty. \quad (4.5)$$

Then, for every $\varepsilon>0$ and $n$ large enough, we have

$$P_r(|R_{q,n}|>\varepsilon)=0.\quad q=1,2$$

**Proof.** Write $R_{q,n}=\frac{1}{n}\sum_{j=2r_np_n+1}^{n}T_{q,j,n}$. As the functions $K_q$, $q=1,2$ are bounded, it follows that

$$|R_{q,n}|\le 2(\frac{n-2r_np_n}{nh_n})\|K_q\|_\infty\le\frac{4p_n\|K_q\|_\infty}{nh_n}.$$

Now $P_r(|R_{q,n}| > \varepsilon) \le P_r(4p_n\|K_q\|_\infty > nh_n\varepsilon)$. According to (4.5) and $h_n \longrightarrow 0$ the lemma follows for $n$ large enough. □

Now, it remains to collect the partial results in order to obtain the exponential rate for kernel estimator centered at its mean.

**Theorem 4.1.** Suppose **A1**, **A2** and (4.5) are satisfied and that

$$-\frac{nh_n^2}{p_n^2}\exp(\frac{nh_n}{p_n})\sum_{j=p_n+2}^{\infty} Cov(X_1, X_j) \le C_1 < \infty. \tag{4.6}$$

Then, for every $\varepsilon \in (0, 6\min(\|K_1\|_\infty, \|K_1\|_\infty^2, \|K_2\|_\infty, \|K_2\|_\infty^2)$ and n large enough,

$$P_r(\frac{1}{n}|\hat{f}_n(x) - E[\hat{f}_n(x)]| > \varepsilon) \le D\exp(-\frac{\varepsilon^2 nh_n^2}{144Cp_n}), \tag{4.7}$$

where $C = \min(\|K_1\|_\infty, \|K_2\|_\infty)$, $D = 2\left[2 + M_1C_1\left(\left(\int K_1'(u)du\right)^2 + \left(\int K_2'(u)du\right)^2\right)\right]$.

**Proof.** We have

$$\hat{f}_n(x) - E[\hat{f}_n(x)] = \left(\hat{f}_{1,n}(x) - E[\hat{f}_{1,n}(x)]\right) + \left(\hat{f}_{2,n}(x) - E[\hat{f}_{2,n}(x)]\right)$$

$$= \frac{1}{n}(Z_{1,n,od} + Z_{1,n,ev}) + \frac{1}{n}(Z_{2,n,od} + Z_{2,n,ev}) + R_{1,n} + R_{2,n}.$$

According to Lemma 4.2 for n large enough, $P\left(|R_{1,n}| > \frac{\varepsilon}{6}\right) = P\left(|R_{2,n}| > \frac{\varepsilon}{6}\right) = 0$. So we need to concentrate only on the first terms. By applying Lemma 4.1 we must check that (4.1) is verified. For this purpose, according to Lemma 2.1, we have

$$|Cov(T_{q,1,n}, T_{q,j,n})| = \frac{1}{h_n^2}\left|Cov(K_q\left(\frac{x - X_1}{h_n}\right), K_q\left(\frac{x - X_j}{h_n}\right))\right| \le \frac{1}{h_n^2}\mathring{M}_1\left(\int K_q'(u)du\right)^2 |Cov^{1/3}(X_1, X_j)|.$$

As $K_1'$ and $K_2'$ are assumed integrable, it follows that (4.6) implies (4.1). Applying then Lemma 4.1, we find for $q = 1,2$

$$P_r(\frac{1}{n}|Z_{q,n,od}| > \frac{\varepsilon}{6}) \le \left\{1 + M_1C_1\left(\int K_q'(u)du\right)^2\right\}\exp\left(-\frac{\varepsilon^2 nh_n^2}{144\, p_n\|K_q\|_\infty}\right),$$

and similarly for $Z_{q,n,ev}$, from where the result follows. □

Now, we use a decomposition of a compact interval to derive a uniform exponential rate for the centered estimator.

**Theorem 4.2.** Suppose **A1**, **A2** and (4.6) are satisfied, the kernel $K$ is Lipschtzian and $p_n = nh_n^3 \longrightarrow +\infty$. Then, for every $\varepsilon \in (0, 6\min(\|K_1\|_\infty, \|K_1\|_\infty^2, \|K_2\|_\infty, \|K_2\|_\infty^2)$, n large enough and each interval $[a, b]$,

$$P_r\left\{\sup_{x\in[a,b]}\left|\widehat{f}_n(x)-E[\widehat{f}_n(x)]\right|>\varepsilon\right\}\le D\frac{b-a}{2}h_n^{-3}\exp\left(-\frac{\varepsilon^2}{576Ch_n}\right),\tag{4.8}$$

where $C$ and $D$ are defined as in Theorem 4.1.

**Proof.** Let $[a,b]$ be a fixed interval and decompose $[a,b]=\bigcup_{j=1}^{s_n}[z_{n,j}-t_n,z_{n,j}+t_n]$ into $s_n$ intervals of length $2t_n$. Then, obviously,

$$\sup_{x\in[a,b]}\left|\widehat{f}_n(x)-E[\widehat{f}_n(x)]\right|\le\max_{1\le j\le s_n}\left|\widehat{f}_n(z_{n,j})-E[\widehat{f}_n(z_{n,j})]\right|$$
$$+\max_{1\le j\le s_n}\sup_{x\in[z_{n,j}-t_n,z_{n,j}+t_n]}\left|\widehat{f}_n(x)-\widehat{f}_n(z_{n,j})-E[\widehat{f}_n(x)-\widehat{f}_n(z_{n,j})]\right|.$$

If we suppose the kernel $K$ to be Lipschitzian, it follows that there exits a constant $\theta>0$ such that

$$\left|\widehat{f}_n(x)-\widehat{f}_n(z_{n,j})-E[\widehat{f}_n(x)-\widehat{f}_n(z_{n,j})]\right|\le2\frac{\theta|x-z_{n,j}|}{h_n^2}\le\frac{2\theta t_n}{h_n^2}.$$

A correct choice of the sequence of radii will verify $\frac{t_n}{h_n^2}\longrightarrow0$. Supposing this condition satisfied, it follows then that

$$P_r\left\{\sup_{x\in[a,b]}\left|\widehat{f}_n(x)-E[\widehat{f}_n(x)]\right|>\varepsilon\right\}\le P_r\left\{\max_{1\le j\le s_n}\left|\widehat{f}_n(z_{n,j})-E[\widehat{f}_n(z_{n,j})]\right|>\varepsilon-\frac{2\theta t_n}{h_n^2}\right\}$$
$$\le s_n\max_{1\le j\le s_n}P_r\left\{\left|\widehat{f}_n(z_{n,j})-E[\widehat{f}_n(z_{n,j})]\right|>\frac{\varepsilon}{2}\right\}.$$

Thus, under the assumptions of Theorem 4.1, as the upper bound derived is independent of $x$, for $K$ Lipschitzian and sequences $s_n$ and $t_n$ such that $b-a=2s_nt_n$ and $\frac{t_n}{h_n^2}\longrightarrow0$, it follows

$$P_r\left\{\sup_{x\in[a,b]}\left|\widehat{f}_n(x)-E[\widehat{f}_n(x)]\right|>\varepsilon\right\}\le Ds_n\exp\left(-\frac{\varepsilon^2nh_n^2}{576Cp_n}\right),\tag{4.9}$$

where $C$ and $D$ are defined as in Theorem 4.1.

Now, choose $t_n=h_n^3$ to which corresponds $s_n=\frac{b-a}{2h_n^3}$ and use (4.9). It is easy to check that (4.5) holds and that this choice for the sequence $t_n$ verifies all the assumptions made. □

Note that there are other possible choices for the sequences. Making these explicit would mean some more precise expressions for $h_n$ and $p_n$. These will be referred in the next section.

## 5. Some examples

In the preceding, we derived some sufficient conditions in order to prove an exponential rate for the kernel estimator for the density. We will now verify that these conditions are not void by constructing examples of covariance structures and choices of the sequences $h_n$ and $p_n$ (which determines $r_n$) that is verified the two assumptions involving these quantities: (4.5) and (4.6). In these examples, we see that there is a tradeoff between the covariance structure and the bandwidth.

**Example 5.1.** Suppose that covariances increase geometrically, namely $Cov(X_1, X_n) = \rho_0\rho^n$, for some $0<\rho<1$, and $\rho_0 < 0$. Then

$$\sum_{j=p_n+2}^{\infty} Cov^{1/3}(X_1, X_j) = \rho_0^{1/3}\frac{\rho^{\frac{p_n+2}{3}}}{1-\rho^{1/3}},$$

so that (4.6) becomes

$$-\frac{nh_n^2\rho_0^{1/3}}{p_n^2(1-\rho^{1/3})}\exp\left\{\frac{nh_n}{p_n}+\frac{p_n+2}{3}\log\rho\right\}\le C_1. \qquad (5.1)$$

**Theorem 5.1.** Suppose **A1**, **A2** and (4.5) are satisfied and $Cov(X_1, X_n) = \rho_0\rho^n$, for some $0<\rho<1$, and $\rho_0<0$. If $\sup_{n\in N}\frac{nh_n}{p_n^2}\le B<\infty$ and $\rho\in(0, e^{-3B})$, then inequality (4.7) holds.

**Proof.** The exponent in (5.1) should be bounded, which is equivalent to $\log\rho \le \frac{3A}{p_n} - 3\frac{nh_n}{p_n^2}$, for some $A\in\Re$. As $p_n \longrightarrow +\infty$ and $\frac{nh_n}{p_n^2}$ is bounded, it is enough that $\log\rho\le -3B$. Finally, note that $\frac{nh_n^2}{p_n^2}\le Bh_n \longrightarrow 0$, so it is bounded. □

**Example 5.2.** Suppose that the covariances increase at the polynomial rate, that is $Cov(X_1, X_n) = a_0 n^{-a}$, for some $a>3$ and $a_0<0$. Then

$$\sum_{j=p_n+2}^{\infty} Cov^{1/3}(X_1, X_j) \sim a_0^{1/3}(p_n+2)^{\frac{3-a}{3}}$$

Inserting this into (4.6) we find a term that behaves like

$$-\frac{nh_n^2}{p_n^2}a_0^{1/3}\exp\left\{\frac{nh_n}{p_n}-(\frac{a-3}{3})\log p_n\right\},$$

as $\frac{\log(p_n+2)}{\log p_n}\sim 1$. If this term is to be bounded, we may have, for some $A>0$

$$h_n \le \frac{Ap_n}{n}+\frac{a-3}{3}\frac{p_n}{n}\log p_n. \qquad (5.2)$$

**Theorem 5.2.** Suppose **A1**, **A2** and (4.5) are satisfied, $Cov(X_1, X_n) = a_0 n^{-a}$, for some $a>3$, $a_0<0$, and $\frac{p_n}{n}\longrightarrow 0$. If $\sup_{n\in N}\frac{nh_n}{p_n\log p_n}\le B<\infty$ and $a>3B+3$, then inequality (4.7) holds.

**Proof.** From the assumptions made it follows easily that $h_n \leq B\frac{p_n}{n}\log p_n$, so (5.2) holds. On the other hand, under these assumptions $\frac{nh_n^2}{p_n^2} \leq \frac{Ah_n}{p_n} + \frac{a-3}{3}\frac{h_n}{p_n}\log p_n \longrightarrow 0$, for some $A>0$, so $\frac{nh_n^2}{p_n^2}$ is bounded. □

## References


[1] Alam, K., Saxena, K.M.L., *Positive dependent in multivariate distributions*, Commun. Statist-Theor. Math. A 10 (1981) 1183-1196.

[2] Bagai, I. and Prakasa Rao, B.L.S., *Kernel-type density and failure rate estimation for associated sequences*, Ann. Inst. Statist. Math. 47 (1995) 253-266.

[3] Bagai, I. and Prakasa Rao, B. L. S. *Estimation of the survival functions for stationary associated processes*, Statist. Probab. Lett. 12 (1991) 291-385.

[4] Block, H.W., Savits, T. H. and Sharked, M., *Some concepts of negative dependence*, Ann. Probab. 10 (1982) 765-772.

[5] Bosq, D., *Nonparametric Statistics for Stochastic Processes*, Lecture Notes in Statistics 110 (1996) Springer, Berlin.

[6] Devroye, L., *Exponential inequalities in nonparametric estimation*, pp. 31-44. In: Nonparametric Functional Estimation and Related Topics, G. Roussas, ed., Kluwer Academic Publishers, Dordrecht.

[7] Henriques, C. and Oliveira, P.E., *Exponential rates for kernel density estimation under association*, Statistica Neerlandica 59(4) (2005) 448-466.

[8] Joag-Dev, K. and Proschan, F., *Negative association of random variables, with application*, Ann. Statist. 11 (1983) 286-295.

[9] Newman, C.M., *Normal fluctuations and the FKG inequalities*, Comm. Math Phys. 74 (1980) 119–128.

[10] Roussas, G.G., *Kernel estimates under association: strong uniform consistency*, Statist. Probab. Lett. 12 (1991) 393-403.

[11] Sadivoka, S.M., *Two-dimensional analogies of an inequality of Esseen with applications to the central limit theorem*, Theory Probab. Appl. 11 (1966) 325–335.