



Asymptotic Confidence Bands for Density and Regression Functions in the Gaussian Case

Nahima Nemouchi and Zaher Mohdeb

Department of Mathematics, Mentouri University, Constantine, Algeria

Received 7 July 2010; Accepted 2 November 2010

Copyright © 2010, Journal Afrika Statistika. All rights reserved

Abstract. In this paper, we obtain asymptotic confidence bands for both the density and regression functions in the framework of nonparametric estimation. Beforehand, the asymptotic behaviors in probability of the kernel estimator of the density and the Nadaraya-Watson estimator of the regression function are described while local and global optimal smoothing parameters are investigated. A simulation study is conducted, showing the good performance of the confidence bands obtained for small sample.

Résumé. Dans cet article, nous obtenons les bandes de confiance asymptotiques pour la densité et la fonction de régression dans le cadre de l'estimation non paramétrique. L'étude a porté également sur le comportement asymptotique en probabilité de l'estimateur à noyau de la densité et de l'estimateur de Nadaraya-Watson de la fonction de régression, ainsi que le paramètre de lissage optimal, local et global. Une étude par simulation est menée pour montrer la bonne performance des bandes de confiance obtenues pour une petite taille d'échantillon.

Key words: Confidence bounds; Density estimation; Kernel estimation; Nonparametric estimation; Regression estimation.

AMS 2010 Mathematics Subject Classification : 62F35, 62G20, 60F15.

1. Introduction

Let $(X_1, Y_1), (X_2, Y_2), \dots$, be independent and identically distributed random replicates of the random vector $(X, Y) \in \mathbb{R}^2$. Let us assume that the marginal distributions of X and Y are normal with unknown means μ_X, μ_Y and standard deviations $\sigma_X > 0, \sigma_Y > 0$ respectively, with an unknown linear correlation coefficient ρ . In the sequel, the following notations and assumptions will be adopted.

$I = [a, b]$ and $J = [a', b']$ denote two fixed intervals of \mathbb{R} such that $-\infty < a' < a < b < b' < \infty$. $|I|$ denotes the Lebesgue measure of I .

K is a real kernel weight function satisfying the following conditions:

K.1 For every $x \in \mathbb{R}$, $K(x) = K(-x)$;

K.2 $\int_{\mathbb{R}} K(t)dt = 1$;

K.3 $[t^2 K] = \int_{\mathbb{R}} t^2 K(t)dt < \infty$;

K.4 $[K^2] = \int_{\mathbb{R}} K^2(t)dt < \infty$;

K.5 $K(u) = 0$ for $u \notin [-\frac{\alpha}{2}, \frac{\alpha}{2}]$, where $\alpha \in]0, \infty[$.

For every $u \in \mathbb{R}$, set $\log_{\theta, K}(u) = \log(\theta \vee u \{\int K^2(t)dt\})$, where $A \vee B = \max(A, B)$ and $\theta > 1$ is a specified constant.

The regression function and the conditional variance of Y given $X = x$ are defined for $x \in J$, by

$$r(x) = E(Y|X = x) = \int_{\mathbb{R}} y \frac{f_{X,Y}(x, y)}{f_X(x)} dx = \mu_Y + \rho \frac{\sigma_X}{\sigma_Y} (x - \mu_X)$$

and

$$v^2(x) = Var(Y|X = x) = \int_{\mathbb{R}} [y - E(Y|X = x)]^2 \frac{f_{X,Y}(x, y)}{f_X(x)} dy = (1 - \rho^2) \sigma_Y^2,$$

where $f_{X,Y}(x, y)$ is the joint density of (X, Y) and $f_X(x)$ is the marginal density of the random variable X .

Let $(h_n)_{n \geq 1}$ be a sequence of real numbers such that $h_n \rightarrow 0$ as $n \rightarrow \infty$. Following Rosenblatt [15], Parzen [13], Nadaraya [11] and Watson [21], the kernel estimators of the functions $f_X(x)$, $r(x)$ and $v^2(x)$ are defined by

$$f_{X,n}(x; h_n) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right),$$

$$r_n(x; h_n) = \begin{cases} \frac{\sum_{i=1}^n Y_i K\left(\frac{x - X_i}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)} & \text{if } \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) \neq 0, \\ \frac{1}{n} \sum_{i=1}^n Y_i & \text{if } \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) = 0 \end{cases}$$

and

$$v_n^2(x; h_n) = \begin{cases} \frac{\sum_{i=1}^n [Y_i - r_n(x; h_n)]^2 K\left(\frac{x - X_i}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)} & \text{if } \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) \neq 0, \\ \frac{1}{n} \sum_{i=1}^n \left[Y_i - \frac{1}{n} \sum_{j=1}^n Y_j \right]^2 & \text{if } \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) = 0, \end{cases}$$

respectively.

Introduce now the following centering factors

$$Ef_{X,n}(x; h_n) = \frac{1}{h_n} E \left[K \left(\frac{x - X}{h_n} \right) \right]$$

$$\widehat{E}r_n(x; h_n) = \begin{cases} \frac{E \left[Y K \left(\frac{X - x}{h_n} \right) \right]}{E \left[K \left(\frac{X - x}{h_n} \right) \right]} & \text{if } E \left[K \left(\frac{X - x}{h_n} \right) \right] \neq 0, \\ E(Y) & \text{if } E \left[K \left(\frac{X - x}{h_n} \right) \right] = 0, \end{cases}$$

which are needed hereafter in our study.

We have

$$\lim_{n \rightarrow \infty} \left\{ E[f_{X,n}(x; h_n)] - f_X(x) \right\} = 0.$$

According to Deheuvels and Mason [3], the conditions $h_n \rightarrow 0$ and $nh_n \rightarrow \infty$ are necessary and sufficient to ensure, that for each $x \in I$, we have

$$\widehat{E}[r_n(x; h_n)] - E[r_n(x; h_n)] = O\left(\frac{1}{nh_n}\right), \quad \text{as } n \rightarrow \infty.$$

From the seminal works by Rosenblatt [15], Parzen [13], Nadaraya [11] and Watson [21], nonparametric function estimation has been widely investigated. Theoretical aspects and properties have been described and numerical implementations have been performed. For an overview on the question, we refer to Prakasa Rao [14], Devroye and Györfi [5], Wand and Jones [20], Bosq and Lecoutre [2], Silverman [18], Nadaraya [12], Scott [17] and the references therein.

The behavior of nonparametric estimators depends strongly upon the smoothing parameter. Several procedures have been proposed describing how to choose this parameter. We quote the cross-validation and the plug-in methods related essentially to the mean square error criterion. Asymptotic properties of estimators need also conditions upon the smoothing parameter. The minimal conditions to set are $h_n \rightarrow 0$ and $nh_n \rightarrow \infty$ as $n \rightarrow \infty$. It is well known that these conditions are necessary and sufficient for the pointwise weak consistency of the kernel estimator of f_X and the Nadaraya-Watson estimator of the regression function r . Moreover, these conditions allow to state a rate of uniform convergence of $\widehat{E}(r_n(x; h_n))$ towards $r(x)$, i.e. we have

$$\left\{ \widehat{E}(r_n(x; h_n)) - r(x) \right\} \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Deheuvels and Mason [3] obtained the rate of uniform convergence in probability for the kernel density estimator. Their result is stated in the following terms. If

$$h_n \rightarrow 0 \quad \text{and} \quad \frac{nh_n}{\log n} \rightarrow \infty, \quad \text{as } n \rightarrow \infty,$$

then

$$\sqrt{\frac{nh_n}{2 \log \left(\frac{1}{h_n} \right)}} \sup_{x \in I} \left| \frac{f_{X,n}(x; h_n) - E[f_{X,n}(x; h_n)]}{\sqrt{2[K^2]^2 f_X(x)}} \right| \xrightarrow{P} 1 \text{ as } n \rightarrow \infty,$$

where I is a compact interval in \mathbb{R} . Deheuvels and Mason [3] have also obtained the optimal rate of convergence in probability for the estimate $r_n(x; h_n)$ of the regression function r . These results allow them to construct asymptotic confidence bands for the functions f_X and r .

Consider now the problem related to the optimal choices $h_n(x)$ and h_n of the smoothing parameter with respect to the mean square error and the integrated mean square error criteria respectively. Our aim is to construct the estimates $\widehat{h}_n(x)$ and \widehat{h}_n of the optimal parameter for these criteria in terms of the empirical estimators $\widehat{\mu}_X$, $\widehat{\mu}_Y$, $\widehat{\sigma}_X$, $\widehat{\sigma}_Y$ and $\widehat{\rho}$, such that

$$\frac{\widehat{h}_n(x)}{h_n(x)} \xrightarrow{P} 1 \quad \text{and} \quad \frac{\widehat{h}_n}{h_n} \xrightarrow{P} 1 \quad \text{as } n \rightarrow \infty.$$

Confidence bounds for $f_X(x)$ and $r(x)$ in terms of $f_{X,n}(x; \widehat{h}_n(x))$ and $r_n(x; \widehat{h}_n(x))$ respectively are then deduced. The interest for our particular case is that we use the bandwidth explicitly in these asymptotic confidence bands.

The remainder of the paper is organized as follows. In section 2, we present the results. A Monte Carlo study is conducted in Section 3. Section 4 is devoted to the proofs.

2. Results

2.1. Mean square error of $f_{X,n}(x; h_n)$

In this section, we give the optimal choices h_n of the smoothing parameter in the expression of the kernel estimators of the functions $f_X(x)$ and $r(x)$ with respect to the mean square error and the integrated mean square error. Having replaced the unknown parameters μ_X and σ_X by their empirical estimators $\widehat{\mu}_X$ and $\widehat{\sigma}_X$ respectively, we also give the asymptotic behaviors of the estimates of the functions f_X and r .

According to Wand and Jones [20], the mean square error of $f_{X,n}(x; h_n)$ is given by

$$E[f_{X,n}(x; h_n) - f_X(x)]^2 = \text{Var} f_{X,n}(x; h_n) + [E(f_{X,n}(x; h_n)) - f_X(x)]^2.$$

This quantity is minimized with respect to h_n , by

$$\begin{aligned} H_{n,1}(x) &= n^{-1/5} \left(\frac{f_X(x) [K^2]}{(f_X''(x) [t^2 K])^2} \right)^{1/5} \\ &= n^{-1/5} \sigma_X \left(\frac{\sqrt{2\pi} [K^2] \exp \frac{1}{2} \left(\frac{x - \mu_X}{\sigma_X} \right)^2}{[t^2 K]^2 \left[\left(\frac{x - \mu_X}{\sigma_X} \right)^2 - 1 \right]^2} \right)^{1/5}. \end{aligned}$$

As μ_X and σ_X are unknown parameters, we replace them by their empirical estimators

$$\hat{\mu}_X = \frac{1}{n} \sum_{i=1}^n X_i$$

and

$$\hat{\sigma}_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \hat{\mu}_X)^2,$$

respectively.

Therefore, we obtain

$$\hat{H}_{n,1}(x) = n^{-1/5} \hat{\sigma}_X \left(\frac{\sqrt{2\pi} [K^2] \exp \frac{1}{2} \left(\frac{x - \hat{\mu}_X}{\hat{\sigma}_X} \right)^2}{[t^2 K]^2 \left[\left(\frac{x - \hat{\mu}_X}{\hat{\sigma}_X} \right)^2 - 1 \right]^2} \right)^{1/5} \quad (1)$$

as an estimator of $H_{n,1}(x)$.

Set now,

$$\Theta_1(x) = \sqrt{\frac{\sigma_X}{[K^2]}} \left(\frac{\sqrt{(2\pi)^3} [K^2]^{1/2} \exp \frac{3}{2} \left(\frac{x - \mu_X}{\sigma_X} \right)^2}{[t^2 K] \left| \left(\frac{x - \mu_X}{\sigma_X} \right)^2 - 1 \right|} \right)^{1/5} \quad (2)$$

and

$$\Theta_{n,1}(x) = \sqrt{\frac{\hat{\sigma}_X}{[K^2]}} \left(\frac{\sqrt{(2\pi)^3} [K^2]^{1/2} \exp \frac{3}{2} \left(\frac{x - \hat{\mu}_X}{\hat{\sigma}_X} \right)^2}{[t^2 K] \left| \left(\frac{x - \hat{\mu}_X}{\hat{\sigma}_X} \right)^2 - 1 \right|} \right)^{1/5} \quad (3)$$

an estimator of $\Theta_1(x)$.

Following Deheuvels and Mason [3], one can see that

$$\hat{h}_{n,1}(x) = \hat{H}_{n,1}(x) / \hat{\sigma}_X, \quad h_{n,1}(x) = H_{n,1}(x) / \sigma_X$$

and $\Theta_{n,1}(x)$ satisfy the following statements.

(B.1) For any $\varepsilon > 0$,

$$P \left(\inf_{x \in I} h_{n,1}(x) - \varepsilon n^{-1/5} \leq \inf_{x \in I} \hat{h}_{n,1}(x) \leq \sup_{x \in I} \hat{h}_{n,1}(x) \leq \sup_{x \in I} h_{n,1}(x) + \varepsilon n^{-1/5} \right) \xrightarrow{P} 1,$$

as $n \rightarrow \infty$.

(B.2) For any $\varepsilon > 0$,

$$P \left(\sup_{x \in I} \left| \frac{\hat{h}_{n,1}(x)}{h_n} - \frac{h_{n,1}(x)}{h_n} \right| \geq \varepsilon \right) \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

(Θ.1) For any $\varepsilon > 0$,

$$P \left(\sup_{x \in I} \left| \frac{\Theta_{n,1}(x)}{\Theta_1(x)} - 1 \right| > \varepsilon \right) \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

2.2. Integrated mean square error of $f_{X,n}(x; h_n)$

Wand and Jones [20] stated that the integrated mean square error relative to $f_{X,n}(x; h_n)$ is given by

$$\int_{\mathbb{R}} E[f_{X,n}(x; h_n) - f_X(x)]^2 dx = \frac{1}{n} [t^2 K] + \int_{\mathbb{R}} [f_X''(x) [t^2 K]]^2 dx + o(h_n^2).$$

The minimum of this quantity, with respect to h_n , is

$$h_{n,2} = n^{-1/5} \left(\frac{[K^2]}{[t^2 K]^2 \int_{\mathbb{R}} [f_X''(x)]^2 dx} \right)^{1/5} = n^{-1/5} \sigma_X \left(\frac{8\sqrt{\pi} [K^2]}{3 [t^2 K]^2} \right)^{1/5}.$$

As the parameter σ_X is unknown, we replace it by the empirical estimator $\hat{\sigma}_X$, to obtain

$$\hat{h}_{n,2} = n^{-1/5} \hat{\sigma}_X \left(\frac{8\sqrt{\pi} [K^2]}{3[t^2 K]^2} \right)^{1/5}. \quad (4)$$

Set

$$\Theta_{n,2}(x) = \left\{ \frac{[K^2]}{\sqrt{2\pi} \hat{\sigma}_X} \exp -\frac{1}{2} \left(\frac{x - \hat{\mu}_X}{\hat{\sigma}_X} \right)^2 \right\}^{-1/2}, \quad (5)$$

an estimator of the following quantity

$$\Theta_2(x) = \left\{ \frac{[K^2]}{\sqrt{2\pi} \sigma_X} \exp -\frac{1}{2} \left(\frac{x - \mu_X}{\sigma_X} \right)^2 \right\}^{-1/2}.$$

As in Deheuvels and Mason [3], one can see that $\hat{h}_{n,2}$ satisfies the statement

$$P \left(h_{n,2} - \varepsilon n^{-1/5} < \hat{h}_{n,2} < h_{n,2} + \varepsilon n^{-1/5} \right) \longrightarrow 1, \text{ as } n \rightarrow \infty, \quad \forall \varepsilon > 0.$$

Moreover, we have

$$P \left(\sup_{x \in I} \left| \frac{\Theta_{n,2}(x)}{\Theta_2(x)} - 1 \right| > \varepsilon \right) \longrightarrow 0, \text{ as } n \rightarrow \infty, \quad \forall \varepsilon > 0.$$

2.3. Mean square error of $r_n(x; h_n)$

In a similar way, according to Wand and Jones [20], the value of $h_{n,3}$ that minimizes the mean square error of $r_n(x; h_n)$ given by

$$E[r_n(x; h_n) - r(x)]^2 = \text{Var}[r_n(x; h_n)] + [E(r_n(x; h_n)) - r(x)]^2$$

is

$$H_{n,3}(x) = n^{-1/5} \left(\frac{\frac{v(x)}{f_X(x)} [K^2]}{\left[r''(x) + 2r'(x) \frac{f'_X(x)}{f_X(x)} [t^2 K] \right]^2} \right)^{1/5} = n^{-1/5} \sigma_X \left(\frac{\sqrt{2\pi} [K^2] (1 - \rho^2) \sigma_X^2 \exp \frac{1}{2} \left(\frac{x - \mu_X}{\sigma_X} \right)^2}{4[t^2 K]^2 \rho^2 x^2} \right)^{1/5}.$$

As σ_X , μ_X and ρ are unknown parameters, we replace them by their empirical estimators, i.e. $\hat{\sigma}_X$, $\hat{\mu}_X$ and

$$\hat{\rho} = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \hat{\mu}_X)(Y_i - \hat{\mu}_Y)}{\hat{\sigma}_X \hat{\sigma}_Y}.$$

Therefore, $H_{n,3}(x)$ is estimated by

$$\hat{H}_{n,3}(x) = n^{-1/5} \hat{\sigma}_X \left(\frac{\sqrt{2\pi} [K^2] (1 - \hat{\rho}^2) \hat{\sigma}_X^2 \exp \frac{1}{2} \left(\frac{x - \hat{\mu}_X}{\hat{\sigma}_X} \right)^2}{4[t^2 K]^2 \hat{\rho}^2 x^2} \right)^{1/5}. \quad (6)$$

Set now

$$\Theta_{n,3}(x) = \left(\frac{(1 - \hat{\rho}^2)^{-2} |\hat{\rho} x|^{-1} \exp -\left(\frac{x - \hat{\mu}_X}{\hat{\sigma}_X} \right)^2}{4\pi \hat{\sigma}_X^{3/2} \hat{\sigma}_Y^5 [K^2]^2 [t^2 K]} \right)^{1/5}, \quad (7)$$

as the estimator of the quantity

$$\Theta_3(x) = \left(\frac{(1 - \rho^2)^{-2} |\rho x|^{-1} \exp -\left(\frac{x - \mu_X}{\sigma_X} \right)^2}{4\pi \sigma_X^{3/2} \sigma_Y^5 [K^2]^2 [t^2 K]} \right)^{1/5}.$$

Similarly, as above, one may show that

$$h_{n,3}(x) = H_{n,3}(x)/\sigma_X, \quad \hat{h}_{n,3}(x) = \hat{H}_{n,3}(x)/\hat{\sigma}_X$$

and $\Theta_{n,3}(x)$ satisfy the statements (B.1), (B.2), (Θ .1).

The asymptotic behaviors of the estimates of the functions f_X and r are described in the following theorems.

Theorem 1. Let $\hat{H}_{n,1}(x)$, $\Theta_{n,1}(x)$ be the estimates given in the statements (1) and (3) respectively. Then, we have

$$\left(\frac{n^{4/5} \hat{\sigma}_X}{2 \log_{\theta,K} \left(\frac{|I|}{n^{-1/5} \hat{\sigma}_X} \right)} \right)^{1/2} \sup_{x \in I} \pm \Theta_{n,1}(x) \left[f_{X,n} \left(x, \hat{H}_{n,1}(x) \right) - f_X(x) \right] \xrightarrow{\mathbf{P}} 1,$$

as $n \rightarrow \infty$.

Remark 1. Note from Theorem 1 that, for any $0 < \varepsilon < 1$, as $n \rightarrow \infty$, we have for any $x \in I$,

$$\lim_{n \rightarrow \infty} P \left(f_X(x) \in \left[f_{X,n}(x, \hat{H}_{n,1}(x)) - (1 + \varepsilon) \Delta_{n,1}(x), f_{X,n}(x, \hat{H}_{n,1}(x)) + (1 + \varepsilon) \Delta_{n,1}(x) \right] \right) = 1$$

and

$$\lim_{n \rightarrow \infty} P \left(f_X(x) \in \left[f_{X,n}(x, \hat{H}_{n,1}(x)) - (1 - \varepsilon) \Delta_{n,1}(x), f_{X,n}(x, \hat{H}_{n,1}(x)) + (1 - \varepsilon) \Delta_{n,1}(x) \right] \right) = 0,$$

where

$$\Delta_{n,1}(x) = \frac{1}{\Theta_{n,1}(x)} \left(\frac{2 \log_{\theta,K} \left(\frac{|I|}{n^{-1/5} \hat{\sigma}_X} \right)}{n^{4/5} \hat{\sigma}_X} \right)^{1/2}.$$

Therefore, the interval

$$\left[f_{X,n}(x, \hat{H}_{n,1}(x)) - \Delta_{n,1}(x), f_{X,n}(x, \hat{H}_{n,1}(x)) + \Delta_{n,1}(x) \right]$$

stands as an asymptotic confidence domain for $f_X(x)$.

Theorem 2. Let $\hat{h}_{n,2}$ and $\Theta_{n,2}(x)$ be the estimates given in the statements (4) and (5) respectively. Then, we have

$$\left(\frac{n \hat{h}_{n,2}}{2 \log_{\theta,K} \left(\frac{|I|}{n^{-1/5} \hat{\sigma}_X} \right)} \right)^{1/2} \sup_{x \in I} \pm \Theta_{n,2}(x) \left[f_{X,n}(x, \hat{h}_{n,2}) - f_X(x) \right] \xrightarrow{\mathbf{P}} 1,$$

as $n \rightarrow \infty$.

Theorem 3. Let $\hat{H}_{n,3}(x)$, $\Theta_{n,3}(x)$ be the estimates given in the statements (6) and (7) respectively. Then, we have

$$\left(\frac{n^{4/5} \hat{\sigma}_X}{2 \log_{\theta,K} \left(\frac{|I|}{n^{-1/5} \hat{\sigma}_X} \right)} \right)^{1/2} \sup_{x \in I} \pm \Theta_{n,3}(x) \left[r_n(x, \hat{H}_{n,3}(x)) - r(x) \right] \xrightarrow{\mathbf{P}} 1$$

as $n \rightarrow \infty$.

Remark 2. Note from Theorem 3 that, for any $0 < \varepsilon < 1$, we have for any $x \in I$,

$$\lim_{n \rightarrow \infty} P \left(r(x) \in \left[r_n(x; \hat{H}_{n,3}(x)) - (1 + \varepsilon) \Delta_{n,2}(x), r_n(x; \hat{H}_{n,3}(x)) + (1 + \varepsilon) \Delta_{n,2}(x) \right] \right) = 1$$

and

$$\lim_{n \rightarrow \infty} P \left(r(x) \in \left[r_n(x; \hat{H}_{n,3}(x)) - (1 - \varepsilon) \Delta_{n,2}(x), r_n(x; \hat{H}_{n,3}(x)) + (1 - \varepsilon) \Delta_{n,2}(x) \right] \right) = 0,$$

where

$$\Delta_{n,2}(x) = \frac{1}{\Theta_{n,3}(x)} \left(\frac{2 \log_{\theta,K} \left(\frac{|I|}{n^{-1/5} \hat{\sigma}_X} \right)}{n^{4/5} \hat{\sigma}_X} \right)^{1/2}.$$

Therefore, we obtain for any $x \in I$, the following asymptotic confidence bounds for the regression function

$$\left[r_n(x; \hat{H}_{n,3}(x)) - \Delta_{n,2}(x), r_n(x; \hat{H}_{n,3}(x)) + \Delta_{n,2}(x) \right].$$

3. Monte Carlo Study

In order to investigate the small sample properties of the confidence bands for the density $f_X(x)$ and regression function $r(x)$, some simulation studies are conducted. We conduct a Monte Carlo study for sample sizes $n = 10$, $n = 50$, $n = 100$ and $n = 800$. We consider the model $Y = X + Z$, with X and Z normally distributed $\mathcal{N}(0, 1)$ and $\mathcal{N}(0, 2)$ respectively. We choose $h_n = n^{(-1/5)}$, $K(x) = 1_{[-1/2, 1/2]}$.

Figure 1 and Figure 2 show the estimators of the density function f_X and the regression function respectively, with the confidence bands.

Examination of Figure 1 and Figure 2 reveal that, when the sample size n increases, the confidence bands become smaller with the density, the regression and their estimators are always in bounds. It appears that the confidence bands have a good performance even for small sample size.

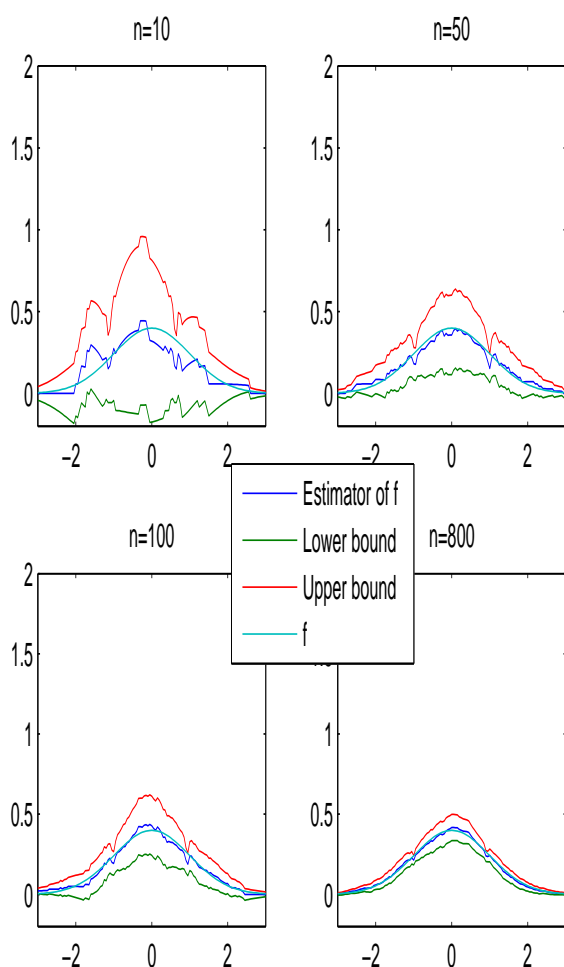


Figure 1. Confidence bands of the density function.

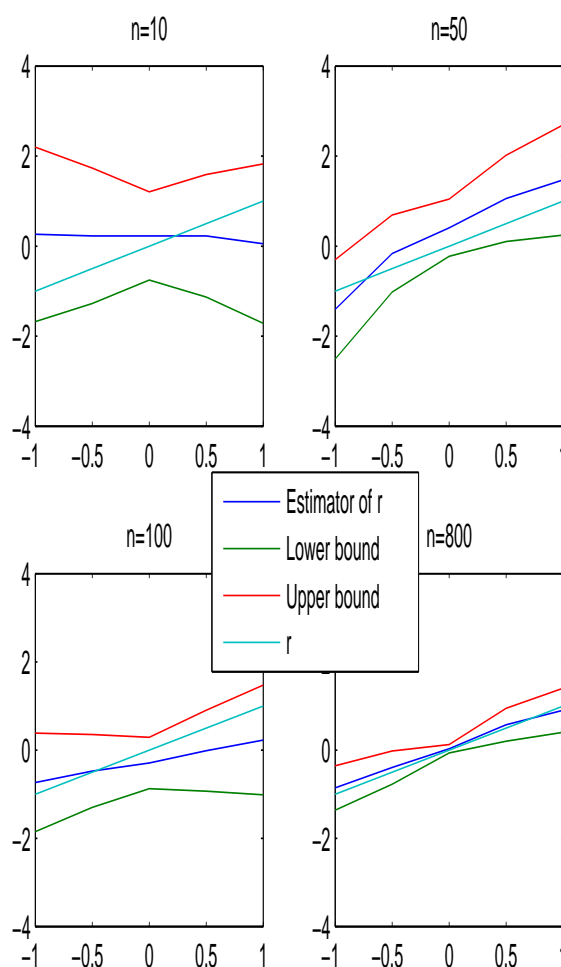


Figure 2. Confidence bands of the regression function

4. Proofs of results

4.1. Proof of Theorem 1

Taking $h_n = n^{-1/5}$, it follows that the conditions (B.1), (B.2) and (Θ_1) are satisfied, Theorem 1.2 of Deheuvels and Mason implies

$$\left(\frac{nh_n}{2 \log_{\theta, K} \left(\frac{|I|}{h_n} \right)} \right)^{1/2} \sup_{x \in \mathbf{I}} \pm \Theta_{n,1}(x) \left[Ef_{X,n}(x, \hat{h}_{n,1}(x)) - f_{X,n}(x, \hat{h}_{n,1}(x)) \right] \\ \xrightarrow{\mathbf{P}} \left\{ \sup_{x \in \mathbf{I}} \frac{\Theta_1^2(x) f_X(x)}{C_1(x)} [K^2] \right\}^{1/2},$$

as $n \rightarrow \infty$, where

$$C_1(x) = \left\{ \frac{\sqrt{2\pi} [K^2] \exp \frac{1}{2} \left(\frac{x - \mu_X}{\sigma_X} \right)^2}{[t^2 K]^2 \left[\left(\frac{x - \mu_X}{\sigma_X} \right)^2 - 1 \right]^2} \right\}^{1/5}. \quad (8)$$

Taking into account the expressions $\Theta_1(x)$ and $C_1(x)$ given by (2) and (8) respectively, we obtain

$$\left(\frac{nh_n}{2 \log_{\theta, K} \left(\frac{|I|}{h_n} \right)} \right)^{1/2} \sup_{x \in \mathbf{I}} \pm \Theta_{n,1}(x) \left[Ef_{X,n}(x, \hat{h}_{n,1}(x)) - f_{X,n}(x, \hat{h}_{n,1}(x)) \right] \xrightarrow{\mathbf{P}} 1, \quad (9)$$

as $n \rightarrow \infty$.

Now, with this choice of the parameter h_n , we have (see Nadaraya (1989))

$$\left(\frac{nh_n}{2 \log_{\theta, K} \left(\frac{|I|}{h_n} \right)} \right)^{1/2} \sup_{x \in \mathbf{I}} \pm \Theta_{n,1}(x) \left[Ef_{X,n}(x, h_n) - f_X(x) \right] \xrightarrow{\mathbf{P}} 0,$$

as $n \rightarrow \infty$.

Let $\hat{h}_n = \hat{\sigma}_X h_n$, we have

$$\frac{\hat{h}_n}{h_n} \xrightarrow{\mathbf{P}} \sigma_X.$$

According to Section 2 of Deheuvels and Mason [3] and works of Einmahl and Mason [8], the locally adapted bandwidth defined by

$$\hat{H}_{n,1}(x) = \hat{h}_n C_1(x) = \hat{\sigma}_X \hat{h}_{n,1}(x)$$

fulfills, via (9)

$$\left(\frac{n^{4/5} \hat{\sigma}_X}{2 \log_{\theta, K} \left(\frac{|I|}{n^{-1/5} \hat{\sigma}_X} \right)} \right)^{1/2} \sup_{x \in \mathbf{I}} \pm \Theta_{n,1}(x) \left[f_{X,n}(x, \hat{H}_{n,1}(x)) - f_X(x) \right] \xrightarrow{\mathbf{P}} 1$$

as $n \rightarrow \infty$.

4.2. Proof of Theorem 2

This proof is based on Corollary 2.1 of Deheuvels and Mason [3], then

$$\left(\frac{n \hat{h}_{n,2}}{2 \log_{\theta, K} \left(\frac{|I|}{\hat{h}_{n,2}} \right)} \right)^{1/2} \sup_{x \in \mathbf{I}} \pm \Theta_{n,2}(x) \left[f_{X,n}(x, \hat{h}_{n,2}) - f_X(x) \right] \\ \xrightarrow{\mathbf{P}} \left\{ \sup_{x \in \mathbf{I}} \pm \Theta_2(x) f_X(x) \int_{\mathbb{R}} K^2(t) dt \right\}^{\frac{1}{2}}$$

i.e.

$$\left(\frac{n \hat{h}_{n,2}}{2 \log_{\theta, K} \left(\frac{|I|}{\hat{h}_{n,2}} \right)} \right)^{1/2} \sup_{x \in \mathbf{I}} \Theta_{n,2}(x) \left[f_{X,n}(x, \hat{h}_{n,2}) - f_X(x) \right] \xrightarrow{\mathbf{P}} 1,$$

as $n \rightarrow \infty$.

4.3. Proof of Theorem 3

According to the Theorem 1.2 and to Section 2 of Deheuvels and Mason [3] and using the works of Einmahl and Mason [8], the proof of this theorem is the same as the proof of Theorem 1.

Acknowledgment

The authors would like to thank the Editors and the anonymous reviewers for their comments and suggestions that improved the quality of the paper.

References

- [1] Akaike, H., 1954. An approximation of the density function. *Ann. Inst. Statist. Math.*, **6**, 127-132.
- [2] Bosq, D. and Lecoutre, J. P., 1987. **Théorie de l'Estimation Fonctionnelle**. Economica, Paris.
- [3] Deheuvels, P. and Mason, D. M. 2004. General asymptotic confidence bands based on kernel-type function estimators. *Stat. Inference Stoch. Process.*, **7**(3), 225-277.
- [4] Devroye, L. 1978. The uniform convergence of the Nadaraya-Watson regression function Estimate. *Can. J. Statist.*, **6**, 179-191.
- [5] Devroye, L. and Györfi, L., 1985. **Nonparametric Density Estimation: The view**. Wiley, New York.
- [6] Devroye, L. and Lugosi, G., 2001. **Combinatorial methods in density estimation**. Springer Series in Statistics. Springer-Verlag, New York, pp. xii+208. ISBN: 0-387-95117-2.
- [7] Einmahl, U. and Mason, D.M., 2000. An empirical process approach to the uniform consistency of kernel-type function estimators. *J. Theoret. Probab.*, **13**(1), 1-37.
- [8] Einmahl, U. and Mason, D.M., 2005. Uniform in bandwidth consistency of kernel-type function estimators. *Ann. Statist.*, **33**(3), 1380-1403.
- [9] Härdle, W., Jansen, P. and Serfling, R., 1988. Strong uniform consistency rates for estimators of conditional functionals. *Ann. Statist.*, **16**(4), 1428-1449.
- [10] Izenman, A.J., 1991. Recent developments in nonparametric density estimation. *J. Amer. Statist. Assoc.*, **86**(413), 205-224.
- [11] Nadaraya, E.A., 1976. On the nonparametric estimator of Bayesian risk in the classification problem. *Proc. AN. Georg SSR*, **82**(2), 277-280 (in Russian).
- [12] Nadaraya, E.A., 1989. Nonparametric estimation of probability densities and regression curves. Translated from the Russian by Samuel Kotz. Mathematics and its Applications (Soviet Series), 20. Kluwer Academic Publishers Group, Dordrecht. pp. x+213. ISBN: 90-277-2757-0.
- [13] Parzen, E., 1962. On estimation of a probability density function and mode. *Ann. Math. Statist.*, **33**, 1065-1076.
- [14] Prakasa, Rao, B.L.S., 1983. *Nonparametric Functional Estimation*. Academic Press, New York.
- [15] Rosenblatt, M., 1956. Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*, **27**, 832-837.
- [16] Roussas, G., 1990. Nonparametric Functional Estimation and Related Topics. NATO ASI series 355. Kluwer, Dordrecht.
- [17] Scott, D.W., 1992. **Multivariate Density Estimation-Theory, Practice and Visualization**. Wiley, New York. ISBN: 978-0-471-54770-9.
- [18] Silverman, B.W., 1986. **Density Estimation for Statistics and Data Analysis**. Chapman & Hall, London, pp. x+175. ISBN: 0-412-24620-1.
- [19] Stone, C.J., 1984. An asymptotically optimal window selection rule for kernel estimates. *Ann. Statist.*, Volume **12**(4), 1285-1297.
- [20] Wand, M.P. and Jones, M.C., 1995. **Kernel Smoothing**. Chapman and Hall. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. ISBN: 0412552701.
- [21] Watson, G.S., 1964. Smooth Regression Analysis. *Sankhya A.*, **26**, 359-372.
- [22] Weissman, I., 1978. Estimation of parameters and large quantiles based on the k largest observations. *J. Amer. Statist. Assoc.* **73**(364), 812-815.
- [23] Woodroffe, M., 1970. On choosing a Delta-Sequence. *Ann. Math. Statist.*, **41**, 1665-1671.