# ABOUT CHOICE OF THE WINDOW WIDTH IN THE KERNEL NONPARAMETRIC ESTIMATE OF PROBABILITY DENSITY

Absava R. M.

Department of Mechanics and Mathematics
Tbilisi State University
380043 University Street 2, Tbilisi, Georgia

*(Received: September 10, 1999)*

*Abstract*

In the present paper the question about finding the window width in the Rozenblatt-Parzen's estimate is considered. The constructed estimate depends only on the choice. Its optimization is proved in the $L_2$ metrics.

*Key words and phrases*: Density function; nonparametric estimate mean integrated square error; kernel estimate.

*AMS subject classification*: 62G05.

1. Let $X_1, X_2, ...., X_n, X_i = (X_i^{(1)}, ..., X_i^{(p)}), i = \overline{1,n}$, be independent observations of the random vector $X = (X^{(1)}, ..., X^{(p)})$ from $R_p$ with the unknown density $f(x), x = (x_1, ....x_p)$.

Let $f(x)$belong to the space $L_2(R_p)$ of all functions square integrable with respect to the Lebesgue measure.

Consider the statistics [1]

$$f_n(x, a_n) = \frac{a_n^p}{n} \sum_{i=1}^{n} K(a_n(x - X_i)), \tag{1}$$

as the estimate of $f(x)$with respect to the given $n$ observations, where $K(x), x \in R_p$, satisfy the conditions formulated below, and $a_n \to \infty$ is a sequence of positive numbers. The value $a_n$ is called window width.

The estimate (1) contains two parameters $K$ and $a_n$, which are to be chosen in some optimal way. It is known that the optimal (in the sense of an asymptotic mean square error of estimate) kernel has the form

$$K_p(x) = \begin{cases} \frac{1}{2} C_p^{-1} (p+2)(1 - x^T \cdot x), & x^T \cdot x \leq 1, \\ 0, & x^T \cdot x > 1. \end{cases}$$

where $C_p$ is the volume of unit $\rho$-dimensional sphere.

If $p = 1$, i.e. $C_1 = 2$, then $K_1(x)$ is the Epanechnikov kernel [2] (see also [1]).

The expression of the optimal window width $a_n^0$[1], obtained with the help of minimization of the asymptotic expression of the mean value of the integral from the square error (m.i.s.e):

$$U(a_n) = E\int (f_n(x, a_n) - f(x))^2 dx$$

is also known (here and later on $\int \equiv \int\limits_{R_n}$). It contains some a priori data that are not always known to the statistician. Below we give a method of the sampling of the window width realizable with sampling (i.e. without any knowledge of the a priori data obtained for which the estimate is asymptotically equivalent to optimal. In other words, on the basis of sampling $X_1, ....X_n$ there are obtained estimates $\{\hat{a}_n\}$ of the elements of the optimal sequence $\{a_n^0\}$ for which m.i.s.e. of the obtained estimate is equivalent (at $n \to \infty$) of m.i.s.e. of the estimate for the optimal sequence $\{a_n^0\}$.

2. In the monograph of E. N. Nadaraya [1] the asymptotic expression $U(a_n)$ is given without a proof.

In this article there is given a method of proof of Theorem 1.2 and Lemma 1.1 from [1] that are more effective than the method of proof of the analogous theorem and lemma developed for one-dimensional case in the mentioned monograph.

Let

$$K(x) = \prod_{j=1}^{p} K_j(x_j), x = (x_1, ...x_p),$$

where $K_j \in H_s$.

$$H_s = \{\varphi : \varphi(-t) = \varphi(t), \quad t \in R_1, \int \varphi(t)dt = 1,$$

$$\sup_{t\in R_1} |\varphi(t)| < \infty, \quad \int t^i \varphi(t)dt = 0, \quad i = \overline{1, s-1},$$

$\int t^s \varphi(t)dt \neq 0$, $\int t^s |\varphi(t)dt| < \infty$, $s \geq 2$ is an even number .

The family of the functions [3]

$$K_c(x) = \begin{cases} 3/8 \left(\frac{c}{5}\right)^{1/2} \left[(3-c) + (3c-5)\frac{cx^2}{5}\right], & |x| \leq \sqrt{\frac{5}{c}}, \\ 0, & |x| > \sqrt{\frac{5}{c}}. \end{cases}$$

where $c \in [1,3]$ belongs to $H_2$. The Bartlett's function [4] $K(x) = \frac{9}{8}(1 - \frac{5}{8}x^2)(= o)$, $|x| \leq 1(|x| > 1)$ belongs to $H_4$ respectively for, and the Nadaraya's function [1]

$$K(x) = \frac{15}{8}\left(1 - \frac{2}{3}x^2 + \frac{1}{15}x^4\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

is a function from the class $H_6$, etc.

We will assume that $f(x), x \in R_p$, has all partial derivatives up to $s$-th$(s \geq 2)$ order inclusively. In addition all partial $s$-th order derivatives are continuous, limited and belong to $L_2(R_p)$. These assumptions will be denoted by $W_s^{(p)}$.

**Theorem 1** ([1]). *If $f(x) \in W_s^{(p)}$ and $K_j(x) \in H_s, j = \overline{1,p}$, then*

$$U(a_n) = \frac{a_n}{n}\int\prod_1^p K_j^2(u)du + a_n^{-2s}\frac{1}{(s!)^2}\int\left(\sum_{j=1}^p \alpha_j \frac{\partial^s f}{\partial x_j^s}\right)^2 dx + o\left(\frac{a_n^p}{n} + a_n^{-2s}\right), \tag{2}$$

*at $n \to \infty$, where,*

$$\alpha_j = \int x^s K_j(x)dx, \quad j = \overline{1,p}.$$

**Proof**. We have

$$U(a_n) = \int Df_n(x,a_n)dx + \int \left[Ef_n(x,a_n)dx - f(x)\right]^2 dx,$$

where

$$\int Df_n(x,a_n)dx = \frac{a_n^{2p}}{n}\int E\prod_{j=1}^p K_j^2\left(a_n\left(x_j - X_1^{(j)}\right)\right) -$$

$$-\frac{1}{n}\int\left(a_n^p E\prod_{j=1}^p K_j\left(x_j - X_1^{(j)}\right)\right)^2 =$$

$$= D_n - E_n.$$

According to Fubini theorem

$$D_n = \frac{a_n^p}{n}\int\left[a_n^p\int\prod_{j=1}^p K_j^2\left(a_n\left(x_j - u_j\right)\right)f(u)du\right]dx = \frac{a_n^p}{n}\int\prod_{j=1}^p K_j^2(u)du.$$

It follows from generalized Minkovskii inequality, that

$$|E_n| \le \frac{1}{n}\int\left(\int\left(\prod_{j=1}^{p}K_j(t_j)f\left(x_j-\frac{t_j}{a_n}\right)dx\right)^2\right)^{1/2}dt=$$

$$=\frac{1}{n}\int f^2(u)du\left(\prod_{j=1}^{p}\int|K_j(t_j)|\,dt_j\right)^{1/2}.$$

So

$$\int Df_n(x,a_n)dx=\frac{a_n^p}{n}\int\prod_{j=1}^{p}K_j^2(u)du+o\left(\frac{a_n^p}{n}\right). \qquad (3)$$

Later, by Taylor formula with the remained term in the integral form we have

$$Ef_n(x,a_n)=\int\prod_{j=1}^{p}K_j(t_j)f\left(x_1-\frac{t_1}{a_n},...,x_p-\frac{t_p}{a_n}\right)dt_1...dt_p=$$

$$=\int\prod_{j=1}^{p}K_j(t_j)\left[\sum_{|l|\le s-1}\frac{t^l}{a_nl}\frac{1}{l!}f^l(x)+R_s(t)\right]dt_1...dt_p,$$

where

$$R_s(t)=s\sum_{|l|=s}\frac{1}{l!}\left(\frac{t}{a_n}\right)^l\int\limits_0^1(1-u)^{l-1}f^{(l)}\left(x+u\frac{t}{a_n}\right)du,$$

$$|l|=\sum_{j=1}^{p}l_j,\quad l!=l_1!...l_p!,$$

$$f^{(l)}(x)=\frac{\partial^{|l|}f(x)}{\partial x_1^{l_1}...\partial x_p^{l_p}},\quad t^l=t_1^{l_1}...t_p^{l_p}.$$

Since $f(x)\in W_s^{(p)}$ and $K_j(x)\in H_s$, $j=\overline{1,p}$, we can write

$$Ef_n(x,a_n)-f(x)=$$

$$=\int\int\limits_0^1\sum_{|l|=s}\frac{1}{l!}\frac{t^p}{a_n^l}\prod_{j=1}^{p}K_j(t_j)(1-u)^{l-1}f^{(l)}\left(x-u\frac{t}{a_n}\right)dudt.$$

Therefore

$$\int (Ef_n(x,a_n) - f(x))^2 \, dx = \frac{s^2}{a_n^{2s}} \int\limits_{\Delta} \int\limits_{\Delta} \Phi_n(t,v,u_1,u_2) dt dv du_1 du_2, \quad (4)$$

where

$$\Phi_n(t,v,u_1,u_2) = \sum_{|l_1|=s} \sum_{|l_2|=s} \frac{t^{l_1}}{l_1!} \frac{v^{l_2}}{l_2!} K(t) K(V) (1-u_1)^{l_1-1} (1-u_2)^{l_2-1} \times$$

$$\times \left( \int f^{(l_1)}(x) f^{(l_2)} \left( x + u_2 \frac{V}{a_n} - \frac{u_1 t}{a_n} \right) dx \right), \qquad \Delta = R_p \times [0,1]. \quad (5)$$

We have further

$$\left| \int f^{(l_1)}(x) f^{(l_2)} \left( x + u_2 \frac{v}{a_n} - \frac{u_1 t}{a_n} \right) dx - \int f^{(l_1)}(x) f^{(l_2)} dx \right| \leq$$

$$\leq \left( \int \left( f^{(l_1)}(x) \right)^2 dx \right)^{1/2} \left( \int \left( f^{(l_2)} \left( x + u_2 \frac{v}{a_n} - \frac{u_1 t}{a_n} \right) - f^{(l_2)}(x) \right)^2 dx \right)^{1/2}. \quad (6)$$

It is well known that any function from $L_2$ is continuous in $L_2$, hence for every fixed $u, \nu, u_1, u_2$ holds

$$\left\| f^{(l_2)} \left( x + u_2 \frac{v}{a_n} - u_1 \frac{t}{a_n} \right) - f^{(l_2)}(x) \right\|_{L_2} \to 0$$

at $n \to \infty$.

It follows from this and (6)

$$\int f^{(l_1)}(x) f^{(l_2)} \left( x + u_2 \frac{v}{a_n} - u_1 \frac{t}{a_n} \right) dx \to \int f^{(l_1)}(x) f^{(l_2)}(x) dx. \quad (7)$$

Moreover

$$|\Phi_n(t,v,u_1,u_2)| \leq \sum_{|l_1|=s} \sum_{|l_2|=s} \frac{|t|^{l_1}}{l_1!} \frac{|v|^{l_2}}{l_2!} |K(t)| \times$$

$$\times |K(v)| \left( \int \left( f^{(l_1)}(x) \right)^2 dx \int \left( f^{(l_2)}(x) \right)^2 dx \right)^{1/2}.$$

These facts permit to apply Lebesque theorem about of the majorized convergence in (5). Thus we obtain

$$\int (Ef_n(x,a_n) - f(x))^2 \, dx = a_n^{-2s} \frac{1}{(s!)^2} \int \left( \sum_{j=1}^{p} \alpha_j \frac{\partial^s f}{\partial x_j^s} \right)^2 dx + o\left(a_n^{-2s}\right).$$

So

$$U(a_n) = \frac{a_n^p}{n} \int \prod_{j=1}^{p} K_j^2(u) du +$$

$$+ a_n^{-2} \frac{1}{(s!)^2} \int \left( \sum_{j=1}^{p} \alpha_j \frac{\partial^s f}{\partial x_j^s} \right)^2 dx + o\left( \frac{a_n^p}{n} + a_n^{-2s} \right).$$

The theorem is proved. □

**Corollary 1** *If $f(x) \in W_2^{(2)}$ and $K_1(x) = K_2(x) = K_0(x) = \frac{1}{2}(= 0)$ respectively for $|x| < 1(|x| \geq 1)$, then from the Theorem 1 the result of G. M. Mania [5] follows .*

Now we will define the optimal value $a_n = a_n^0$, minimizing the asymptotic expression (at $n \to \infty$ of m.i.s.e. $U(a_n)$.

**Lemma 1** ([6]) . *Let $A, B, \alpha$ and $\beta$ be given positive numbers. Then*

$$\min_{x>0} \left( Ax^{\alpha} + Bx^{-\beta} \right) = (\alpha + \beta) \left\{ \left( \frac{A}{\beta} \right)^{\beta} \left( \frac{B}{\alpha} \right)^{\alpha} \right\}^{\frac{1}{\alpha+\beta}}$$

*and the minimum is reached for the value of $x$*

$$x_{\min} = \left( \frac{\beta B}{\alpha A} \right)^{\frac{1}{\alpha+\beta}}.$$

Assume

$$A = \frac{1}{n} \int \prod_{j=1}^{p} K_j^2(u) du, \quad B = \frac{1}{(s!)^2} \int \left( \sum_{j=1}^{p} \alpha_j \frac{\partial^s f}{\partial x_j^s} \right)^2 dx,$$

$$\alpha = p, \quad \beta = 2s.$$

Then from Lemma 1 we obtain

$$a_n^0 = \theta n^{-\gamma}, J = \frac{1}{2s+p},$$

$$\theta^{2s+p} = 2s \int \left( \sum_{j=1}^{p} \alpha_j \frac{\partial^s f}{\partial x_j^s} \right)^2 dx \left( (s!)^2 p \int \prod_{j=1}^{p} K_j^2(u) du \right)^{-1}. \quad (8)$$

If we substitute the obtained optimal value $a_n^0$ in the right hand of (2) we will have

$$U(a_n^0) = R(s, f, k) n^{-\frac{2s}{2s+p}} + o\left( n^{-\frac{2s}{2s+p}} \right),$$

where

$$R(s,f,k) = (2s+p)\left\{\left((2s)^{-1}\int\prod_{j=1}^{p}K_j^2(u)du\right)^{2s}\times\right.$$

$$\left.\times\left(p^{-1}(s!)^{-2}\int\left(\sum_{j=1}^{p}\alpha_j\frac{\partial^s f}{\partial x_j^s}\right)^2 dx\right)^p\right\}^{\frac{1}{2s+p}}.$$

Thus the optimal estimate of the density $f_n(x,a_n)$ is the integral consistent of the order $N = n^{2s/(2s+p)}$, i.e. $N\cdot U(a_n^0)\to 0$ to the finite nonlinear limit at $n\to\infty$.

**Lemma 2** *If $f(x)\in W_s^{(p)}$ and $K_j(x)\in H_s,\ j=\overline{1,p}$, where $K_j(x),\ j=\overline{1,p}$ have continuous partial derivatives up to $s\geq 2$ order inclusively, $K_j(x)\to 0,\ i=\overline{1,s-1},\ j=\overline{1,p}$ as $x\to\pm\infty$ and $\int x^s\left|K_j^{(s)}(x)\right|dx<\infty,\ j=\overline{1,p}$ then $n\to\infty$.*

$$\int\left[\sum_{j=1}^{p}\alpha_j E\frac{\partial^s f_n(x,a_n)}{\partial x_j^s}\right]^2 dx\to\int\left[\sum_{j=1}^{p}\alpha_j\frac{\partial^s f_n(x,a_n)}{\partial x_j^s}\right]^2 dx$$

*at $n\to\infty$.*

Lemma 2 is generalization of the corresponding Lemma 1.1 of Nadaraya [1].

**Proof**. From the identity

$$a_n^{p+s}K_m^{(s)}(x_m-u_m)\prod_{j\neq 1}^{p}K_j\left((x_j-u_j)\,a_n\right)f(u_1,...,u_p)=$$

$$=a_n^p\prod_{j=1}^{p}K_j\left(a_n\left(x_j-u_j\right)\right)\frac{\partial^s f(x)}{\partial x_m^s}-$$

$$-\frac{\partial}{\partial u_m}\left(\sum_{j=1}^{s-1}a_n^{j+p}K_m^j\left(a_m(x_m-u_m)\right)\frac{\partial^{(s-j-1)}f(x)}{\partial x_m^{s-j-1}}\prod_{\substack{j=1\\ j\neq m}}^{p}K_j\left(a_n\left(x_j-u_j\right)\right)\right),$$

taking into account that $K_j^{(s)}(x)\to 0$ at $|x|\to\infty,\ s=0,1,\ j=\overline{1,p}$ it follows

$$E\frac{\partial^s f_n(x,a_n)}{\partial x_m^s}=\int a_n^{p+s}K_m^{(s)}\left((x_m-u_m)a_n\right)\times$$

$$\times\prod_{\substack{j=1\\ j\neq m}}^{p}K_j\left(a_m\left(x_j-u_j\right)\right)f(u_1,...,u_p)du_1,...du_p= \tag{9}$$

$$=\int\prod_{j=1}^{n}K_j(t_j)f_m^{(s)}\left(x+\frac{t}{a_n}\right)dt,$$

where

$$f_m^{(s)}(t)=\frac{\partial^s f(t)}{\partial t_m^s},\quad t=(t_1,...,t_p).$$

Further, we have

$$\int\left(\sum_{j=1}^{p}\alpha_j E\frac{\partial^s f_n(x,a_n)}{\partial x_j^s}\right)^2dx=$$

$$=\sum_{j=1}^{p}\alpha_j^2\int\left(E\frac{\partial^s f_n}{\partial x_j^s}\right)^2dx+\sum_{i\neq j}\alpha_i\alpha_j\int E\frac{\partial^s f_n}{\partial x_i^s}E\frac{\partial^s f_n}{\partial x_j^s}dx.$$

Thus, it remains only to prove that

$$\int E\frac{\partial^s f_n}{\partial x_i^s}E\frac{\partial^s f_n}{\partial x_j^s}dx\to\int f_i^{(s)}(x)f_j^{(s)}(x)dx$$

at $n\to\infty$.

It follows from (9)

$$\int E\frac{\partial^s f_n}{\partial x_i^s}E\frac{\partial^s f_n}{\partial x_j^s}dx=\int\int K(t)K(u)dtdu\int f_i^{(s)}(z)f_j^{(s)}\left(z+\frac{t-u}{a_n}\right)dz, \tag{10}$$

where $K(t)=\prod_{j=1}^{p}K_j(t_j)$.

Since, analogous to (7)

$$\int f_i^{(s)}(z)f_j^{(s)}\left(z+\frac{t-u}{a_n}\right)dz\to\int f_i^{(s)}(z)f_j^{(s)}(z)dz$$

and

$$\left|K(t)K(u)\int f_i^{(s)}(z)f_j^{(s)}(z+\frac{t-u}{a_n})dz\right|\leq$$

$$\leq |K(t)|\,|K(u)| \left( \int \left( f_i^{(s)}(z) \right)^2 dz \int \left( f_j^{(s)}(z) \right)^2 dz \right)^{1/2},$$

from (10) and Lebesque theorem we obtain

$$\int E \frac{\partial^s f_n}{\partial x_i^s} E \frac{\partial^s f_n}{\partial x_j^s} dx \to \int f_i^{(s)}(z) f_j^{(s)}(z) dz.$$

The Lemma is proved. □

We will use the following

**Lemma 3** ([1]). *Let random variables have absolute moments up to the m-th order, in addition to probability unit* $\eta_n \geq d_n > 0$, $I_n \geq 2d_n > 0$ *and* $d_n \to 0$ *at* $n \to \infty$. *If* $E\,|\eta_n - I_n|^m = O(a_n^m)$, *where* $a_n^k = O(d_n)$ *for some* $k > 0$, *then* $E\,|\eta_n^\alpha - I_n^\alpha|^m = O(a_n^m)$, *where* $-\alpha_0 < \alpha \leq 1,\ \alpha_0 > 0$.

3. Now we shall get down to solving the problem formulated at the beginning of the paper. We shall assume that $f(x) \in W_s^{(p)}$ and $K_j(x) \in H_s$ satisfy the conditions of Lemma 2. Consider the obtained optimal value (8), $a_n^0 = \theta n^\gamma, \gamma = \frac{1}{2s+p}$ supplying the minimum of m.i.s.e. . First, since $\theta = \theta(f,k)$ is unknown we will estimate it by sampling $X_1, ..., X_n$. Let $\{t_n\}$ be a sequence of positive numbers such that $t_n \to \infty$ at $n \to \infty$, where $t_n = o(n^\alpha), \alpha = \frac{1}{(2s+p)^2}$. Further, let $\{b_n\}$ be a sequence of positive numbers converging to zero and satisfying the condition

$$nb_n \geq C > 0$$

(here and later $C, C_1, C_2, ...$ will be positive constants).

We shall introduce the notation:

$$f_{ni}^{(s)}(x) = \frac{\partial^s f_n(x, t_n)}{\partial x_j^s}, \quad \mu_{ni}^{(s)}(x) = E_{ni}^{(s)}(x),$$

$$\Phi_p(u) = \sum_{j=1}^{p} \alpha_i K_j^{(s)}(u_j) \prod_{\substack{r=1 \\ r \neq j}}^{p} K_r(u_r).$$

The properties of estimates of $f_n$ and $f_{ni}^{(s)}(x)$ defined by (1) prompt us to consider the sequence of estimates of $\theta^{2s+p}$ of the form

$$\hat{\theta}_n^{2s+p} = l(k,s) \left[ \int \left( \sum_{j=1}^{p} \alpha_j f_{ni}^{(s)}(x) \right)^2 dx + b_n \right], \tag{11}$$

where

$$l(k,s)=\frac{2s}{p(s!)^2}\left[\int\prod_{j=1}^{p}K_j^2(u_j)du_j\right]^{-1}.$$

Let's assume

$$\hat{\theta}_n^{2s+p}=l(k,s)\left[\int\left(\sum_{j=1}\alpha_i\mu_{ni}^{(s)}(x)\right)^2dx+b_n\right],$$

$$\hat{a}_n=\hat{\theta}_n n^{\gamma},\ \sigma=\theta_n n^{\gamma},\ \gamma=\frac{1}{2s+p},\ \tau_n^2=\frac{t_n^{2(2s+p)}}{n}=o(n^{-(2s+p-1)\gamma}).$$

It is not difficult to show that

$$\hat{\theta}_n^{2s+p}=l(k,s)\left[\int\Omega_n^2(x)dx+b_n\right],$$
$$\theta_n^{2s+p}=l(k,s)\left[\int\left(E\Omega_n^2(x)\right)^2dx+b_n\right], \tag{12}$$

where

$$\Omega_n(x)=\frac{t_n^{s+p}}{n}\sum_{i=1}^{p}\Phi\left(t_n\left(x-X_i\right)\right), \tag{13}$$

Let

$$\Phi_p^*(u)=\int\Phi_p(v)\Phi_p(u-v)dv,$$
$$T_n(u)=t^pt_n^p\int\Phi_p^*(t_n\left(u-V\right))f(V)dV. \tag{14}$$

Then by the definition of $\Omega_n(x)$ we obtain the correlation

$$ET_n(x_1)=t_n^{-2s}\int\left(E\Omega_n(x)\right)^2dx. \tag{15}$$

**Lemma 4** *If $f(x)$ and $K_j(x)$, $j=\overline{1,p}$ satisfy the conditions of Lemma 2, then*

$$\left|\hat{\theta}_n-\theta_n\right|^m=O(\tau_n^m), \tag{16}$$

*where $m>0$ is an integer number.*

**Proof**. Using (13), (14) and (15), we obtain

$$E\left[\int\left(\Omega_n(x)\right)^2dx-\int\left(E\Omega_n(x)\right)^2dx\right]^{2m}=$$

$$= t_n^{4sm} E\left[n^{-2}\sum_{j=1}^{n}\sum_{i=1}^{n} t_n^p \Phi_p^*\left(t_n(X_j - X_i)\right) - ET_n(x_1)\right]^{2m}.$$

From this, according to the inequality

$$\left|\sum_{k=1}^{m} a_k\right|^r \le m^{r-1}\sum_{k=1}^{m} |a_k|^2,$$

where $r \ge 1$ is an integer number, we find

$$E\left[\int (\Omega_n(x))^2\, dx - \int (E\Omega_n(x))^2\, dx\right]^{2m} \le C_1 t_n^{4ms}\left(E_n^{(1)} + E_n^{(2)}\right),$$

where

$$E_n^{(1)} = E\left[n^{-2}\sum_{j=1}^{n}\sum_{i=1}^{n} t_n^p \Phi_p^*\left(t_n(X_j - X_i)\right) - n^{-1}\sum_{i=1}^{n} T_n(x_i)\right]^{2m},$$

$$E_n^{(2)} = E\left[n^{-1}\sum_{j=1}^{n}\left(T_n(X_j) - ET_n(X_j)\right)\right]^{2}.$$

Let us estimate each of $E_n^{(1)}$ and $E_n^{(2)}$ individually.

Taking into account the easily verifying inequalities

$$\left|\Phi_p^*(u)\right| \le C_2$$
$$|T_n(u)| \le C_3 t_n^p \qquad (17)$$

we have

$$E_n^{(1)} \leq C_4 \left( \frac{t_n^p}{n} \right)^{2m} +$$

$$+n^{-1} E \sum_{j=1}^{n} \left| \left[ n^{-1} \sum_{1 \leq i \neq j \leq n} t_n^p \Phi_p^* \left( t_p (X_j - X_i) \right) - T(X_j) \right] \right|^{2m} =$$

$$= C_4 \left( \frac{t_n^p}{n} \right)^{2m} + E \left[ n^{-1} \sum_{i=2}^{n} t_n^p \Phi_p^* \left( t_n (X_i - X_1) \right) - T_n(X_1) \right]^{2m} \leq$$

$$\leq C_5 \left( \frac{t_n^p}{n} \right)^{2m} + E \left( E \left( \frac{1}{n-1} \sum_{i=2}^{n} t_n^p \Phi_p^* \left( t_n (X_i - X_1) \right) - T_n(X_1) \right)^{2n} / X_1 \right) =$$

$$= C_5 \left( \frac{t_n^p}{n} \right)^{2m} + E_n^{(3)}, \tag{18}$$

where

$$E_n^{(3)} = E \left( E \left( \frac{1}{n-1} \sum_{i=2}^{n} t_n^p \Phi_p^* \left( t_n (X_i - X_1) \right) - T_n(X_1) \right)^{2m} / X_1 \right).$$

It is clear that $t_n^p \Phi_p^* \left( t_n (X_i - X_1) \right) - T_n(X_1)$ are independent identically distributed random variables for the given $X_1$. In addition

$$E \left( t_n^p \Phi_p^* \left( t_n (X_i - X_1) \right) - T_n(X_1) / X_1 \right) = 0.$$

□

Therefore for the estimate of $E_n^{(3)}$ we can use the theorem of Petrov [7] relating to the estimate of the moments of the sum of independent random variables.

**Theorem 2** (Petrov V.V.([7])). *Let $X_1, X_2, ..., X_n$ be independent random variables, $EX_k = 0,\ k = \overline{1,n}\ , p \geq 2$. Then*

$$E \left| \sum_{k=2}^{n} X_k \right|^p \leq C(p) n^{p/2-1} \sum_{k=2}^{n} E \left| X_k \right|^p .$$

By the theorem of Petrov and (17) we obtain

$$E_n^{(3)} = E\left(E\left(\frac{1}{n-1}\sum_{i=2}^{n} t_n^p \Phi_p^*\left(t_n(X_i - X_1)\right) - T_n(X_1)\right)^{2m} / X_1\right) \leq$$

$$\leq C(m)\frac{n^{m-1}}{(n-1)^{2m}}\times$$

$$\times E\left[\sum_{i=2}^{n} E\left|t_n^p \Phi_p^*\left(t_n(X_i - X_1)\right) - T_n(X_1)\right|^{2m} / X_1\right] \leq$$

$$\leq \widetilde{C}(m)\frac{n^{m-1}}{(n-1)^{2m}} t_n^{2mp} n = O\left(\frac{t_n^{2p}}{n}\right)^m.$$

So

$$E_n^{(1)} = O\left(\frac{t_n^{2p}}{n}\right)^m. \tag{19}$$

Analogously we find

$$E_n^{(2)} = O\left(\frac{t_n^{2p}}{n}\right)^m. \tag{20}$$

Further substituting (19) and (20) in (17), we get

$$\left[\int (\Omega_n(x))^2\, dx - \int (E\Omega_n(x))^2\, dx\right]^{2m} =$$

$$= O\left(t_n^{4sm}\frac{t_n^{2m}}{n^m}\right) = O\left(\frac{t_n^{2(2s+p)}}{n}\right)^n = O\left(\tau_n^2\right)^m. \tag{21}$$

Hence by definitions of $\hat{\theta}_n^{2s+p}$ and $\theta_n^{2s+p}$ we obtain from (21)

$$E\left|\hat{\theta}_n^{2s+p} - \theta_n^{2s+p}\right|^{2m} = O\left(\tau_n^{2m}\right). \tag{22}$$

From this, in particular, we have

$$E\left|\hat{\theta}_n^{2s+p} - \theta_n^{2s+p}\right|^{2m} = \left(\tau_n^m\right). \tag{23}$$

Thus, since, $\hat{\theta}_n^{2s+p} \geq l(k,s)b_n, \tau_n^4/b_n \leq C_6\frac{t_n^{4(2s+p)}}{n} \leq C_6\frac{1}{n^{1-2/(2s+p)}} \leq C_7$, i.e. $\tau_n^4 = O(b_n)$ and by virtue of Lemma 2 $\theta_n^{2s+p} \geq \frac{\theta_n^{2s+p}}{2}$, for $n > N$, then according to (23) and Lemma 3 we obtain (16). □

**Corollary 2** *$\hat{\theta}_n$ is a consistent estimate for $\theta$.*

**Theorem 3** *Let $f(x)$ and $K_j(x), j=\overline{1,p}$ satisfy the conditions of Lemma 2 and in addition let the function*

$$K_1(x) = pK(x) + \sum_{j=1}^{p} x_i K_j^{(1)}(X_j) \prod_{\substack{r=1 \\ r\neq j}}^{p} K_r(x_r)$$

*admit a nondiecreasing and integrable majorant $K_0(x), K_0(\pm x) = K_0(x)$, in the interval $R_+^p = [0,\infty)^p$. Then*

$$U(\hat{a}_n) \sim U(a_n^0), \tag{24}$$

*at $n \to \infty$ (relation $\alpha_n \sim \beta_n$ means that, $\frac{\alpha_n}{\beta_n} \to 1$).*

Theorem 3 is generalization of Theorem 1.3 of Nadaraya [1].

**Proof**. It follows from (2) and lemma 2 that $\theta_n \to \theta$ and $V(\sigma_n) \sim U(a_n^0)$, where $\sigma_n = \theta_n n^\gamma,\ a_n^0 = \theta n^\gamma,\ \gamma = \frac{1}{2s+p}$. From the representation

$$\frac{U(\hat{a}_n)}{U(\sigma_n)} = 1 + 2\frac{E\int (f_n(x,\hat{a}_n) - f_n(x,\sigma_n))(f_n(x,\sigma_n) - f(x))\,dx}{U(\sigma_n)}$$

$$+\frac{E\int (f_n(x,\hat{a}_n) - f_n(x,\sigma_n))^2\,dx}{U(\sigma_n)}$$

and the Cauchy-Schwarz inequality it follows that it is sufficient to show

$$E\int (f_n(x,\hat{a}_n) - f_n(x,\sigma_n))^2\,dx = o(n^{-2s\gamma}) \tag{25}$$

for the proof of (24).

By the finite increment formula and Cauchy-Schwarz inequality we get

$$E\int (f_n(x,\hat{a}_n) - f_n(x,\sigma_n))^2\,dx \leq$$

$$\leq E^{1/2}\left[\int \left(\frac{\partial f_n(x,u)}{\partial u}\right)^2_{u=\xi_n} dx\right]^2 E^{1/2}\left(\hat{\theta}_n - .\theta_n\right)^4 n^{2\gamma},$$

where random variable $\xi_n$ lies between $\hat{a}_n$ and $\sigma_n : |\sigma_n - \xi_n| \leq |\hat{a}_n - \sigma_n|$.

But by (16) and $\tau_n^2 = o(n^{-(2s+p-1)\gamma})$, we have

$$E^{1/2}\left(\hat{\theta}_n - \theta_n\right)^4 = O(\tau_n^2) = o(n^{-(2s+p-1)\gamma}),\ p \geq 1.$$

Therefore it remains to show that

$$n^{4\gamma}E\left[\int\left(\frac{\partial f_n(x,u)}{\partial u}\right)^2_{u=\xi_n}dx\right]^2=O(1). \tag{26}$$

Let $A_n$ be events stating that $\hat{\theta}_n>\frac{\theta}{2}$ and let $I_{A_n}$ be the indicator of event $A_n$. Let, further, $\lambda_n=\frac{\theta}{2}n^{\gamma}$. It is easy to see that

$$n^{4\gamma}E\left[\int\left(\frac{\partial f_n(x,u)}{\partial u}\right)^2_{u=\xi_n}dx\right]^2=$$

$$=n^{4\gamma}\left[\int\left(\frac{\xi_n^{p-1}}{n}\sum_{j=1}^{n}K\left(\xi_n\left(x-X_i\right)\right)\right)^2\right]^2=n^{4\gamma}\left(B_n^{(1)}+B_n^{(2)}\right),$$

where

$$B_n^{(1)}=E\left[\int\left(\frac{\xi_n^{p-1}}{n}\sum_{i=1}^{n}K\left(\xi_n\left(x-X_i\right)\right)\right)^2dxI_{A_n}\right]^2,$$

$$B_n^{(2)}=E\left[\int\left(\frac{\xi_n^{p-1}}{n}\sum_{i=1}^{n}K\left(\xi_n\left(x-X_i\right)\right)\right)^2dxI_{\overline{A}_n}\right]^2.$$

We have $\sigma_n\geq\frac{\theta}{2}n^{\gamma}$ and $\hat{\theta}_n\geq\frac{\theta}{2}n^{\gamma}$ on the $\omega$-set $A_n$ for any sufficiently large $n$. Consequently $\xi_n\geq\lambda_n=\frac{\theta}{2}n^{\gamma}$ for $\omega\in A_n$. Therefore, taking into account that $K_1(x)$ has the majorant $K_0(x)$, we find

$$n^{4\gamma}B_n^{(1)}\leq C_8E\left[\frac{\xi_n^{2p}}{\lambda_n^{2p}}\int\left(\frac{\lambda_n^p}{n}\sum_{i=1}^{n}K_0\left(\lambda_n\left(x-X_i\right)\right)\right)^2dx\right]^4\leq$$

$$\leq C_8E^{1/2}\left(\frac{\xi_n}{\lambda_n}\right)^{8p}E^{1/2}\left[\int\left(\frac{\lambda_n^p}{n}\sum_{i=1}^{n}K_0\left(\lambda_n\left(x-X_i\right)\right)\right)^2dx\right]^4.$$

Now, basing on the method of the proof of (21) and on Lemma 2, we conclude

$$E\left[\int\left(\frac{\lambda_n^p}{n}\sum_{i=1}^{n}K_0\left(\lambda_n\left(x-X_i\right)\right)\right)^2dx\right]^4=O(1).$$

On the other hand

$$E\left(\frac{\xi_n}{\lambda_n}\right)^{8p} \le C_9 + 2^{8p} E\left[\frac{\sigma_n - \hat{a}_n}{\lambda}\right]^{8p} = C_9 + C_{10} E\left|\hat{\theta}_n - \theta_n\right|^{8p} = O(1).$$

So

$$n^{4\gamma}\left(B_n^{(1)}\right) = O(1). \qquad (27)$$

Assume now

$$K_0^*(X) = \int K_0(V) K_0(x - V) dV.$$

Then

$$B_n^{(2)} \le E\left[\frac{1}{\xi_n n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} K_0^*\left(\xi_n\left(X_i - X_j\right)\right) I_{\overline{A}_n}\right]^2 \le C_{11} E\left(\frac{1}{\xi_n} I_{\overline{A}_n}\right)^2.$$

From the definition of $\hat{\theta}_n^{2s+p}$ and the fact that $nb_n \ge C > 0$ we obtain

$$\hat{a}_n = \hat{\theta}_n n^{\gamma} \ge \left[l\left(k, s\right)\right]^{\gamma} \left(nb_n\right)^{\gamma} \ge \left(Cl\left(k, s\right)\right)^{\gamma} = C_{12} \ne 0,$$

and $\sigma_n \ge C_{12}$. Therefore, $\xi_n \ge C_{12}$.

Hence

$$n^{4\gamma} B_n^{(2)} \le C_{13} n^{4\gamma} p(\overline{A}_n).$$

But taking into account (16) we have

$$p(\overline{A}_n) \le p\left\{\left|\hat{\theta}_n - \theta\right| \ge \theta_n - \frac{\theta}{2}\right\} \le C_{14} M \left|\hat{\theta}_n - \theta\right|^4 = O(\tau_n^4) =$$

$$= o\left(n^{-(2s+p-1)2\gamma}\right)$$

for any sufficiently large $n$.

Hence

$$n^{4\gamma} B_n^{(2)} = O(1). \qquad (28)$$

Finally, (27) and (28) imply (26). Therefore the theorem is proved. □

Note. It follows from (25) and decomposition of $U(a_n^0)$ that

$$U(\hat{a}_n) = U(a_n^0) + o\left(n^{-\frac{2s}{2s+p}}\right).$$

*References*

1. Nadaraya, E. A. *Nonparametric estimation of probability densities and regression curves.* Kluwer Academic Publishers, Dordrecht, Holland, 1989.

2. Epanechinikov, V. A. *Nonparametric estimate of many-dimensional probability density.* Theory of probabilities and its appl. 14(1969), no.1, 156-161.

3. Ghosh, B. K.; Huang, W. M. *The power and optimal kernel of the Bickel-Rosenblatt test for goodness of fit.* Ann. Stat. 19(1991), no.2, 999-1009.

4. Bartlett, M. *Statistical estimation of density functions.* Sankhya, I. Stat 25(1963), 245-254.

5. Mania, G. M. *Statistical estimate of probability distribution.* Tbilisi Univ. Press, Tbilisi, 1974.

6. Parzen, E. *On estimation of a probability density function and mode.* Ann. Math. Stat. 33(1962), no.3, 1065-1076.

7. Petrov, V.V. *Limitiny theorems for sums of independent random variables.* "Nayka", Moskow, 1987.