Transactions of A. Razmadze
Mathematical Institute
Vol. **175** (2021), issue 3, 319–326

ON THE INTEGRAL SQUARE DEVIATION BETWEEN TWO KERNEL TYPE ESTIMATORS OF THE BERNOULLI REGRESSION FUNCTIONS FOR THE GROUP DATA

PETRE BABILUA

Abstract. In the paper, the limit distribution of an integral square deviation between two kernel type of Nadaraya–Watson estimators of the Bernoulli regression function for the group data is established.

1. Introduction

Let random variables $Y^{(i)}$, $i=1,2$ take two values: 1 and 0 with probabilities p_i ("success") and $1-p_i$ ("failure"). Assume that the probability of "success" p_i is the function of an independent variable $x\in[0,1]$, i.e., $p_i=p_i(x)=\mathbf{P}\{Y^{(i)}=1 \mid \ x\}$ [1,2,6]. Let x_j, $j=1,\dots,n$, be the points of division of the interval $[0,1]$:

$$x_j=\frac{2j-1}{2n}\,,\ \ j=1,\dots,n.$$

Let, further, $Y_{ij}^{(k)}$, $j=1,\dots,m_i^{(k)}$, $i=1,\dots,n$, $k=1,2$, be mutually independent Bernoulli random variables with $\mathbf{P}\{Y_{ij}^{(k)}=1 \mid \ x_i\}=p_k(x_i)$, $\mathbf{P}\{Y_{ij}^{(k)}=0 \mid \ x_i\}=1-p_k(x_i)$, $j=1,\dots,m_i^{(k)}$, $i=1,\dots,n$, $k=1,2$.

Based on the group samplings $Y_{ij}^{(k)}$, $j=1,\dots,m_i^{(k)}$, $i=1,\dots,n$, $k=1,2$, let us introduce the Nadaraya–Watson type estimates for the $p_1(x)$ and $p_2(x)$ Bernoulli regression functions:

$$\widehat{p}_{kn}(x)=p_{kn}(x)\cdot p_n^{-1}(x),$$

$$p_{kn}(x)=\frac{1}{nb_n}\sum_{i=1}^{n}K\Big(\frac{x-x_i}{b_n}\Big)\overline{Y}_i^{(k)},\ \ \overline{Y}_i^{(k)}=\frac{1}{m_i^{(k)}}\sum_{j=1}^{m_i}Y_{ij}^{(k)},$$

$$p_n(x)=\frac{1}{nb_n}\sum_{i=1}^{n}K\Big(\frac{x-x_i}{b_n}\Big),\ \ k=1,2,$$

where $K(x)$ is some distribution density that satisfies the requirements formulated below, $b_n\to 0$ is a sequence of positive integers, and $\widehat{p}_{kn}(x)$ is a kernel estimate of the regression function (see [4,7]).

2. Assumptions and Notation

Assume that the kernel $K(x)\geq 0$ is chosen such that it is a function with finite variation and satisfies the conditions: $K(x)=K(-x)$, $K(x)=0$ for $|x|\geq\tau>0$, $\int K(x)\,dx=1$. We denote the class of such functions by $H(\tau)$.

For comparison of estimates $\widehat{p}_{1n}(x)$ and $\widehat{p}_{2n}(x)$ we introduce the statistics

$$U_n=\frac{1}{2}\,nb_n\int\limits_{\Omega_n}\big[\widetilde{p}_{1n}(x)-\widetilde{p}_{2n}(x)\big]^2\,dx,$$

$$\widetilde{p}_{in}(x)=p_{in}(x)-\mathbf{E}\,p_{in}(x),\ \ i=1,2$$

2020 *Mathematics Subject Classification.* Primary: 62G05, Secondary: 62G08.
Key words and phrases. Bernoulli regression function; Nadaraya–Watson estimator; Limit distribution.

and the following notation:

$$Q_{ij} = \psi_n(x_i, x_j), \quad \psi_n(u,v) = \int\limits_{\Omega_n} K\Big(\frac{x-u}{b_n}\Big) K\Big(\frac{x-v}{b_n}\Big)\,dx,$$

$$B_n^2(p_1,p_2) = (nb_n)^{-2} \sum_{k=2}^{n} \Big[\frac{p_1(x_k)(1-p_1(x_k))}{m_k^{(1)}} + \frac{p_2(x_k)(1-p_2(x_k))}{m_k^{(2)}}\Big] \times \sum_{i=1}^{k-1} \Big[\frac{p_1(x_i)(1-p_1(x_i))}{m_i^{(1)}} + \frac{p_2(x_i)(1-p_2(x_i))}{m_i^{(2)}}\Big] Q_{ik}^2,$$

$$\eta_{ij}^{(n)} = \frac{\varepsilon_i \varepsilon_j Q_{ij}}{nb_n B_n(p_1,p_2)}\,, \quad \varepsilon_i = \varepsilon_{1i} - \varepsilon_{2i}, \ \ \varepsilon_{ki} = \overline{Y}_i^{(k)} - p_k(x_i), \ \ i = 1,\dots,n, \ \ k = 1,2,$$

$$\xi_k^{(n)} = \sum_{i=1}^{k-1} \eta_{ik}^{(n)}, \ \ k = 2,\dots,n, \ \ \xi_1^{(n)} = 0, \ \ \xi_k^{(n)} = 0, \ \ k > n,$$

$$\mathcal{F}_k^{(n)} = \sigma(\varepsilon_1,\dots,\varepsilon_k),$$

where $\mathcal{F}_k^{(n)}$ is a σ-algebra generated by random variables $\varepsilon_1,\dots,\varepsilon_k$, $\mathcal{F}_0^{(n)} = (\varnothing,\Omega)$ (in the sequel, for the sake of simplicity, we will write ξ_k, η_{ij} and $\mathcal{F}_k$ instead of $\xi_k^{(n)}$, $\eta_{ij}^{(n)}$ and $\mathcal{F}_k^{(n)}$).

3. Auxiliary Results

Lemma 1. *A stochastic sequence $(\xi_k,\mathcal{F}_k)_{k\ge1}$ is a martingale difference.*

Lemma 2 ([5])**.** *Let $K(x) \in H(\tau)$ and $p(x)$, $0 \le x \le 1$, be a function of bounded variation. If $nb_n \to \infty$, then*

$$\frac{1}{nb_n} \sum_{i=1}^{n} K^{\nu_1}\Big(\frac{x-x_i}{b_n}\Big) K^{\nu_2}\Big(\frac{y-x_i}{b_n}\Big) p^{\nu_3}(x_i) = \frac{1}{b_n} \int\limits_0^1 K^{\nu_1}\Big(\frac{x-u}{b_n}\Big) K^{\nu_2}\Big(\frac{y-u}{b_n}\Big) p^{\nu_3}(u)\,du + O\Big(\frac{1}{nb_n}\Big),$$

uniformly with respect to $x, y \in [0,1]$, where $\nu_i \in N \cup \{0\}$, $i = 1,2,3$.

Lemma 3. *Let $K(x) \in H(\tau)$ and $p_k(x) \in C^1[0,1]$, $k = 1,2$. If $nb_n^2 \to \infty$, then*

$$b_n^{-1} N_n^2 B_n^2(p_1,p_2) \ge b_n^{-1} \sigma_n^2(p_1,p_2) \longrightarrow \sigma^2(p_1,p_2) \ \ \text{as} \ \ n \to \infty,$$

where

$$\sigma_n^2(p_1,p_2) = (nb_n)^{-2} \sum_{k=2}^{n} d(x_k) \sum_{i=1}^{k-1} d(x_i) Q_{ik}^2, \ \ d(x_i) = \sum_{r=1}^{2} p_r(x_i)(1-p_r(x_i)), \ \ i = 1,\dots,n,$$

$$\sigma^2(p_1,p_2) = \frac{1}{2} \int\limits_0^1 d^2(x)\,dx \int\limits_{|t|\le 2\tau} K_0^2(t)\,dt,$$

$$d(x) = \sum_{r=1}^{2} p_r(x)(1-p_r(x)),$$

$$N_n = \max(N_n^{(1)}, N_n^{(2)}), \quad N_n^{(k)} = \max_{1\le i\le n} m_i^{(k)}, \ \ k = 1,2, \ \ K_0 = K * K.$$

In particular, if $m_i^{(1)} = m_i^{(2)} = N_n$, $i = 1,\dots,n$, then

$$\lim_{n\to\infty} \frac{N_n^2 B_n^2(p_1,p_2)}{b_n} = \sigma^2(p_1,p_2)$$

and also, if $p_1(x) = p_2(x) = p_0(x)$, then

$$\lim_{n\to\infty} \frac{N_0^2 B_n^2(p_1,p_2)}{b_n} = \sigma^2(p_0) = 2 \int p_0^2(x)(1-p_0(x))^2\,dx \int\limits_{|x|\le 2\tau} K_0^2(x)\,dx.$$

Proof. Clearly,

$$N_n^2 b_n^{-1} B_n^2(p_1,p_2) \geq b_n^{-1}\sigma_n^2(p_1,p_2)$$

and

$$\sigma_n^2(p_1,p_2) = A_n^{(1)}(p_1,p_2) + A_n^{(2)}(p_1,p_2),$$

where

$$A_n^{(1)}(p_1,p_2) = \frac{1}{2}\,(nb_n)^{-2}\sum_{i,k=1}^{n} d(x_i)d(x_k)Q_{ik}^2,$$

$$A_n^{(2)}(p_1,p_2) = -\frac{1}{2}\,(nb_n)^{-2}\sum_{i=1}^{n} d^2(x_i)Q_{ii}^2.$$

It is easy to verify that

$$b_n^{-1}|A_n^{(2)}(p_1,p_2)| = \frac{1}{2}\,n^{-2}b_n^{-3}\sum_{i=1}^{n} d^2(x_i)\Big(\int\limits_{\Omega_n} K^2\Big(\frac{x-x_i}{b_n}\Big)\,dx\Big)^2 \leq c_1\,\frac{1}{nb_n}\,. \tag{1}$$

Further, using the definition of Q_{ki}, we obtain

$$A_n^{(1)}(p_1,p_2) = \frac{1}{2}\,(nb_n)^{-2}\int\limits_{\Omega_n}\int\limits_{\Omega_n}\Big(\sum_{i=1}^{n} d(x_i)K\Big(\frac{x-x_i}{b_n}\Big)K\Big(\frac{y-x_i}{b_n}\Big)\Big)^2\,dx\,dy.$$

Since $p(x)\in C^1[0,1]$ and $[\frac{x-1}{b_n}\,,\frac{x}{b_n}]\supset[-\tau,\tau]$ for all $x\in\Omega_n(\tau)$ and by Lemma 2, it is easy to verify that

$$b_n^{-1}A_n^{(1)}(p_1,p_2) = \frac{1}{2}\int\limits_{\Omega_n(\tau)} d^2(x)\,dx\int\limits_{\frac{x-1}{b_n}+\tau}^{\frac{x}{b_n}-\tau} K_0^2\Big(\frac{x-y}{b_n}\Big)\,dx\,dy + O(b_n) + O\Big(\frac{1}{nb_n^2}\Big).$$

Therefore

$$b_n^{-1}A_n^{(1)}(p_1,p_2) \longrightarrow \frac{1}{2}\int\limits_0^1 d^2(x)\,dx\int\limits_{|x|\leq 2\tau} K_0^2(x)\,dx. \tag{2}$$

From the last, in particular, for $p_1(x)=p_2(x)=p_0(x)$, it follows that

$$b_n^{-1}A_n^{(1)}(p_0,p_0) \longrightarrow \frac{1}{2}\int\limits_0^1 p_0^2(x)(1-p_0(x))\,dx\int\limits_{|x|\leq 2\tau} K_0^2(x)\,dx. \tag{3}$$

From (1), (2) and (3) follows the assertion. □

4. Asymptotic Normality of Statistic U_n

The following assertion holds true.

Theorem. *Let $K(x)\in H(\tau)$ and $p(x)\in C^1[0,1]$. If $\frac{N_n^4}{nb_n^2}\to\infty$ and $N_n^4 b_n\to 0$ as $n\to\infty$, then for the hypothesis $H_0:\ p_1(x)=p_2(x)$,*

$$\frac{U_n-\mathbf{E}\,U_n}{B_n}\xrightarrow{d} N(0,1),\quad B_n=B_n(p_1,p_2),$$

where $\xrightarrow{d}$ denotes convergence in distribution, and $N(0,1)$ is a random variable having a standard normal distribution $\Phi(x)$.

Proof. We have

$$\frac{U_n-\mathbf{E}\,U_n}{B_n} = H_n^{(1)} + H_n^{(2)},$$

where

$$H_n^{(1)} = \sum_{k=1}^{n} \xi_k, \quad H_n^{(2)} = \frac{\sum\limits_{i=1}^{n} (\varepsilon_i^2 - \mathbf{E}\,\varepsilon_i^2) Q_{ii}}{n b_n B_n}\,.$$

Let us show that $H_n^{(2)}$ converges to zero in probability. Indeed,

$$\operatorname{Var} H_n^{(2)} \le \frac{1}{(n b_n B_n)^2} \sum_{i=1}^{n} (\mathbf{E}\,\varepsilon_{1i}^4 + \mathbf{E}\,\varepsilon_{2i}^4) Q_{ii}^2$$

Since $Q_{ij} \le c_1 b_n$ and $\mathbf{E}\,\varepsilon_i^4 \le c_2 (m_i^{(1)})^{-2} + c_3 (m_i^{(2)})^{-2}$, by Lemma 3, we have

$$\operatorname{Var} H_n^{(2)} \le c_4 \, \frac{N_n^2}{n b_n^2} \longrightarrow 0.$$

Therefore $H_n^{(2)} \xrightarrow{\mathbf{P}} 0$ (here and below the letter $\mathbf{P}$ over the arrow denotes convergence in probability).

Now let us show that $H_n^{(1)} \xrightarrow{d} N(0,1)$. To this end, we need to verify the validity of Corollaries 2 and 6 of Theorem 2 in [3]. We have to show the fulfilment of the conditions contained in these corollaries and guaranteeing the asymptotic normality of the square integrable martingale difference, which, according to Lemma 1, is our sequence $\{\xi_k, \mathcal{F}_k\}_{k \ge 1}$. It is easy to see that $\sum\limits_{k=1}^{n} \mathbf{E}\,\xi_k^2 = 1$. Asymptotic normality $H_n^{(1)}$ takes place if for $n \to \infty$,

$$\sum_{k=1}^{n} \mathbf{E}\left[\xi_k^2 I\big(|\xi_k| \ge \varepsilon\big) \mid \mathcal{F}_{k-1}\right] \xrightarrow{\mathbf{P}} 0 \tag{4}$$

and

$$\sum_{k=1}^{n} \xi_k^2 \xrightarrow{\mathbf{P}} 1. \tag{5}$$

In [3], it is proved that if (5) and the condition $\sup\limits_{1 \le k \le n} |\xi_k| \xrightarrow{\mathbf{P}} 0$ are fulfilled, then condition (4) is fulfilled, too.

Since for $\varepsilon > 0$,

$$\mathbf{P}\Big\{ \sup_{1 \le k \le n} |\xi_k| \ge \varepsilon \Big\} \le \varepsilon^{-4} \sum_{k=1}^{n} \mathbf{E}\,\xi_k^4,$$

according to relation (7) given below, to prove

$$H_n^{(1)} \xrightarrow{d} N(0,1).$$

It remains only to verify (5). For this, it suffices to ascertain that

$$\mathbf{E}\Big(\sum_{k=1}^{n} \xi_k^2 - 1\Big)^2 \longrightarrow 0 \ \text{ as } \ n \to \infty,$$

i.e., that since $\sum\limits_{k=1}^{n} \mathbf{E}\,\xi_k^2 = 1$,

$$\mathbf{E}\Big(\sum_{k=1}^{n} \xi_k^2\Big)^2 = \sum_{k=1}^{n} \mathbf{E}\,\xi_k^4 + 2 \sum_{1 \le k_1 < k_2 \le n} \mathbf{E}\,\xi_{k_1}^2 \xi_{k_2}^2 \longrightarrow 1. \tag{6}$$

Let us prove (6). Taking the definitions of ξ_k and η_k into account, we write

$$\sum_{k=1}^{n} \mathbf{E}\,\xi_k^4 = I_n^{(1)} + I_n^{(2)},$$

where

$$I_n^{(1)} = \frac{1}{(nb_n)^4\sigma_n^4}\sum_{k=2}^{n}\mathbf{E}\,\varepsilon_k^4\sum_{j=1}^{k-1}\mathbf{E}\,\varepsilon_j^4 Q_{jk}^4,$$

$$I_n^{(2)} = \frac{3}{(nb_n)^4 B_n^4}\sum_{k=2}^{n}\sum_{i\neq j}\mathbf{E}\,\varepsilon_j^2\mathbf{E}\,\varepsilon_i^2 Q_{jk}^2 Q_{ik}^2.$$

Since

$$Q_{ij} \le c_6 b_n,\quad \mathbf{E}\,\varepsilon_j^4 \le c_5,\quad \mathbf{E}\,\varepsilon_j^2 \le c_6,\quad |\mathbf{E}\,\varepsilon_j^3| \le c_7,$$

and

$$b_n^{-1}\sigma_n^2(p_1,p_2) \longrightarrow \sigma^2(p_1,p_2),$$

we have

$$I_n^{(1)} = O\Big(\frac{N_n^4}{(nb_n)^2}\Big),\quad I_n^{(2)} = O\Big(\frac{N_n^4}{nb_n^2}\Big).$$

Therefore

$$\sum_{k=1}^{n}\mathbf{E}\,\xi_k^4 \xrightarrow{\ \mathbf{P}\ } 0 \ \text{ as } \ n\to\infty. \tag{7}$$

Further, from the definition of ξ_i, it follows that

$$\xi_{k_1}^2\xi_{k_2}^2 = B_{k_1k_2}^{(1)} + B_{k_1k_2}^{(2)} + B_{k_1k_2}^{(3)} + B_{k_1k_2}^{(4)},$$

where

$$B_{k_1k_2}^{(1)} = \sigma_2(k_1)\sigma_2(k_2),\quad B_{k_1k_2}^{(2)} = \sigma_2(k_1)\sigma_1(k_2),$$

$$B_{k_1k_2}^{(3)} = \sigma_1(k_1)\sigma_2(k_2),\quad B_{k_1k_2}^{(4)} = \sigma_1(k_1)\sigma_1(k_2),$$

$$\sigma_1(k) = \sum_{i\neq j\neq j\le k-1}\eta_{ik}\eta_{jk},\quad \sigma_2(k) = \sum_{i=1}^{k-1}\eta_{ik}^2.$$

Consequently,

$$2\sum_{1\le k_1<k_2\le n}\mathbf{E}\,\xi_{k_1}^2\xi_{k_2}^2 = \sum_{i=1}^{4}A_n^{(i)},$$

where

$$A_n^{(i)} = 2\sum_{1\le k_1<k_2\le n}\mathbf{E}\,B_{k_1k_2}^{(i)},\quad i=1,2,3,4.$$

Let us consider $A_n^{(3)}$. Using the definition of η_{ij}, it can be easily shown that

$$\mathbf{E}\,B_{k_1k_2}^{(3)} = 0,$$

and therefore

$$A_n^{(3)} = 0. \tag{8}$$

Let us estimate $A_n^{(2)}$. We have

$$\big|\mathbf{E}\,B_{k_1k_2}^{(2)}\big| = \frac{1}{(nb_nB_n)^4}\Big|\sum_{i=1}^{k_1-1}\mathbf{E}\,\varepsilon_i^3\mathbf{E}\,\varepsilon_{k_1}^3\,\mathbf{E}\,\varepsilon_{k_2}^2 Q_{ik_1}^2 Q_{ik_2} Q_{k_1k_2}\Big|.$$

Since $\mathbf{E}\,|\varepsilon_i^3| \le c_7$ and $Q_{ij}\le c_6 b_n$, we obtain

$$\big|\mathbf{E}\,B_{k_1k_2}^{(2)}\big| \le c_8\,\frac{k_1-1}{(nB_n)^4}\,.$$

Further, because

$$\sum_{1\le k_1<k_2\le n}(k_1-1) = O(n^3) \ \text{ and } \ b_n^{-1}\sigma_n^2 \longrightarrow \sigma^2(p) > 0,$$

we have

$$|A_n^{(2)}| \le c_9\,\frac{n^3}{n^4B_n^4} = c_9\,\frac{N_n^4}{nb_n^2(b_n^{-1}N_n^2B_n^2)^2} = O\Big(\frac{N_n^4}{nb_n^2}\Big). \tag{9}$$

Now, we will establish that $A_n^{(1)} \to 1$ as $n\to\infty$. It is obvious that

$$A_n^{(1)} = 2\sum_{1\le k_1<k_2\le n} \mathbf{E}\,B_{k_1k_2}^{(1)} = S_n^{(1)} + S_n^{(2)},$$

where

$$S_n^{(1)} = 2\sum_{1\le k_1<k_2\le n}\Big(\sum_{i=1}^{k_1-1}\mathbf{E}\,\eta_{ik_1}^2\Big)\Big(\sum_{j=1}^{k_2-1}\mathbf{E}\,\eta_{jk_2}^2\Big),$$

$$S_n^{(2)} = 2\Big(\sum_{k_1<k_2}\mathbf{E}\,B_{k_1k_2}^{(1)} - \sum_{k_1<k_2}\Big(\sum_{i=1}^{k_1-1}\mathbf{E}\,\eta_{ik_1}^2\Big)\Big(\sum_{j=1}^{k_2-1}\mathbf{E}\,\eta_{jk_2}^2\Big)\Big).$$

From the definition of σ_n^2 it follows that

$$S_n^{(1)} = 1 - \sum_{k=2}^{n}\Big(\sum_{i=1}^{k-1}\mathbf{E}\,\eta_{ik}^2\Big)^2,$$

moreover,

$$\sum_{k=2}^{n}\Big(\sum_{i=1}^{k-1}\mathbf{E}\,\eta_{ik}^2\Big)^2 \le c_{10}\,\frac{b_n^4n^3}{(nb_n)^4B_n^4} = O\Big(\frac{1}{nb_n^2}\Big).$$

Thus

$$S_n^{(1)} = 1 + O\Big(\frac{N_n^2}{nb_n^2}\Big). \tag{10}$$

We will further show that $S_n^{(2)}\to 0$. It is easy to see that

$$S_n^{(2)} = 2\sum_{k_1<k_2}\Big[\sum_{i=1}^{k_1-1}\operatorname{cov}(\eta_{ik_1}^2,\eta_{ik_2}^2) + \sum_{i=1}^{k_1-1}\operatorname{cov}(\eta_{ik_1}^2,\eta_{k_1k_2}^2)\Big].$$

Therefore

$$\operatorname{cov}(\eta_{ik_1}^2,\eta_{ik_2}^2) = O\Big(\frac{1}{n^4B_n^4}\Big).$$

Next, since

$$\sum_{1\le k_1<k_2\le n}(k_1-1) = O(n^3),$$

we have

$$S_n^{(2)} = O\Big(\frac{1}{nB_n^4}\Big) = O\Big(\frac{N_n^4}{nb_n^2}\Big). \tag{11}$$

Thus by (10) and (11),

$$A_n^{(1)} = 1 + O\Big(\frac{N_n^4}{nb_n^2}\Big). \tag{12}$$

Finally, we will prove that $A_n^{(4)}\to 0$ as $n\to\infty$. Using the definition of η_{ij} and the relations $Q_{ij}\ge 0$ and $\mathbf{E}\,\varepsilon_i^2 \le c_{11}$, we have

$$\mathbf{E}\,B_{k_1k_2}^{(4)} = 4\sum_{1\le t<s\le k_1-1}\mathbf{E}\,\eta_{sk_1}\eta_{tk_1}\eta_{sk_2}\eta_{tk_2} \le \frac{c_8}{n^4b_n^4B_n^4}\sum_{1\le t<s\le k_1-1}Q_{sk_1}Q_{tk_1}Q_{sk_2}Q_{tk_2}.$$

Therefore

$$A_n^{(4)} \le \frac{c_9}{n^2b_n^4B_n^4}\sum_{k_1<k_2}A_{k_1k_2},$$

where

$$A_{k_1k_2} = \frac{1}{n^2}\sum_{1\le t<s\le k_1-1}Q_{sk_1}Q_{tk_1}Q_{sk_2}Q_{tk_2}.$$

But

$$\sum_{k_1<k_2} A_{k_1k_2} \le \sum_{k_1,k_2=1}^{n} \Big(\frac{1}{n}\sum_{t=1}^{n} Q_{tk_1}Q_{tk_2}\Big)^2.$$

Thus

$$A_n^{(4)} \le c_{10}\,\frac{1}{n^2b_n^4B_n^4} \times \sum_{k_1,k_2=1}^{n}\Big(\int\limits_{\Omega_n(\tau)}\int\limits_{\Omega_n(\tau)} K\Big(\frac{x-x_{k_1}}{b_n}\Big)K\Big(\frac{y-x_{k_2}}{b_n}\Big)\cdot\frac{1}{n}\sum_{i=1}^{n}K\Big(\frac{x-x_i}{b_n}\Big)K\Big(\frac{y-x_i}{b_n}\Big)\,dx\,dy\Big)^2. \quad (13)$$

Next, applying Lemma 2, from (13), we conclude that

$$A_n^{(4)} \le \frac{c_{12}}{n^2b_n^4B_n^4} \times \sum_{k_1,k_2=1}^{n}\Big\{\frac{1}{n}\int\limits_0^1\int\limits_{\Omega_n(\tau)}\int\limits_{\Omega_n(\tau)} K\Big(\frac{x-x_{k_1}}{b_n}\Big)\cdot K\Big(\frac{y-x_{k_2}}{b_n}\Big)\cdot K\Big(\frac{x-u}{b_n}\Big)\cdot K\Big(\frac{y-u}{b_n}\Big)\,du\,dx\,dy\Big\}^2 + O\Big(\frac{N_n^4}{nb_n^2}\Big). \quad (14)$$

Further, applying again Lemma 2 in (14), it can be shown that

$$A_n^{(4)} \le \frac{c_{12}}{b_n^4B_n^4}\int\limits_0^1\int\limits_0^1\int\limits_0^1\int\limits_0^1 \psi_n(u_1,v_2)\psi_n(u_1,v_1)\psi_n(u_2,v_1)\psi_n(u_2,v_2)\,du_1\,du_2\,dv_1\,dv_2 + O\Big(\frac{N_n^4}{nb_n^2}\Big), \quad (15)$$

where

$$\psi_n(x,y) = \int\limits_{\Omega_n} K\Big(\frac{t-x}{b_n}\Big)K\Big(\frac{t-y}{b_n}\Big)\,dt.$$

Let us now estimate the integral contained in (15). Since

$$\Big[\frac{x-1}{b_n}, \frac{x}{b_n}\Big] \supseteq [-\tau,\tau] \ \text{ for all } \ x\in\Omega_n(\tau),$$

we have

$$\int\limits_0^1 \psi_n(u_1,v_2)\psi_n(u_1,v_1)\,du_1 = b_n\int\limits_{\overline{\Omega}_n(\tau)} K\Big(\frac{t-v_2}{b_n}\Big)K\Big(\frac{z-v_1}{b_n}\Big)K_2\Big(\frac{z-t}{b_n}\Big)\,dt\,dz \le c_{13}b_n^3,$$

$$K_2 = K*K, \quad \overline{\Omega}_n(\tau) = \Omega_n(\tau)\times\Omega_n(\tau).$$

Therefore

$$A_n^{(4)} \le c_{14}\,\frac{1}{b_nB_n^4}\int\limits_0^1\int\limits_0^1\int\limits_0^1 \psi_n(u_2,v_1)\psi_n(u_2,v_2)\,du_2\,dv_1\,dv_2 + O\Big(\frac{N_n^4}{nb_n^2}\Big). \quad (16)$$

Applying the same operations to (16), we finally obtain

$$A_n^{(4)} \le c_{15}\,\frac{b_n^4}{b_nB_n^4} + O\Big(\frac{N_n^4}{nb_n^2}\Big) = O\Big(\frac{b_n^3N_n^4}{b_n^2(\frac{N_n^2B_n^2}{b_n})^2}\Big) + O\Big(\frac{N_n^4}{nb_n^2}\Big) = O(b_nN_n^4) + O\Big(\frac{N_n^4}{nb_n^2}\Big). \quad (17)$$

After combining relations (8), (9), (12) and (17), we establish that

$$2\sum_{1\le k_1<k_2\le n} \mathbf{E}\,\xi_{k_1}^2\xi_{k_2}^2 \longrightarrow 1.$$

From this and (7), it follows that

$$\mathbf{E}\Big(\sum_{k=1}^{n}\xi_k^2-1\Big)^2 \longrightarrow 0 \ \text{ as } \ n\to\infty.$$

Thus

$$\frac{U_n-\mathbf{E}\,U_n}{B_n}\xrightarrow{d} N(0,1).$$

The theorem is proved. □

Corollary. *Let* $K(x)\in H(\tau)$, $p_k(x)\in C^1[0,1]$, $k=1,2$. *Further, let,* $m_i^{(1)}=m_i^{(2)}=N_n$, $i=1,\dots,n$. *If* $\frac{N_n^4}{nb_n^2}\to 0$ *and* $N_n^4 b_n\to 0$, *then*

$$b_n^{-\frac{1}{2}}\,\frac{N_nU_n-\mathbf{E}\,U_n}{\sigma(p_1,p_2)}\xrightarrow{d} N(0,1).$$

References

1. J. B. Copas, Plotting p against x. *Appl. Statist.* **32** (1983), no. 2, 25–31.
2. S. Efromovich, *Nonparametric Curve Estimation.* Methods, theory, and applications. Springer Series in Statistics. Springer-Verlag, New York, 1999.
3. R. Sh. Lipcer, A. N. Shirjaev, A functional central limit theorem for semimartingales. (Russian) *Teor. Veroyatnost. i Primenen.* **25** (1980), no. 4, 683–703.
4. E. A. Nadaraya, On a regression estimate. (Russian) *Teor. Verojatnost. i Primenen.* **9** (1964), 157–159.
5. E. Nadaraya, P. Babilua, G. Sokhadze, Estimation of a distribution function by an indirect sample. *Ukr. Mat. Zh.* **62** (2010), no. 12, 1642–1658 and *Ukr. Math. J.* **62** (2010), no. 12, 1906–1924.
6. H. Okumura, K. Naito, Weighted kernel estimators in nonparametric binomial regression. The International Conference on Recent Trends and Directions in Nonparametric Statistics. *J. Nonparametr. Stat.* **16** (2004), no. 1-2, 39–62.
7. G. S. Watson, Smooth regression analysis. *Sankhya Ser. A* **26** (1964), 359–372.

(Received 27.07.2021)

Ivane Javakhishvili Tbilisi State University, Faculty of Exact and Natural Sciences, Department of Mathematics, 13 University Str., Tbilisi 0186, Georgia

E-mail address: petre.babilua@tsu.ge