Transactions of A. Razmadze
Mathematical Institute
Vol. **174** (2020), issue 2, 135–146

ON THE TESTING HYPOTHESIS OF EQUALITY OF TWO BERNOULLI REGRESSION FUNCTIONS

PETRE BABILUA AND ELIZBAR NADARAYA

Abstract. We establish the limit distribution of the square-integrable deviation of two nonparametric kernel-type estimations for the Bernoulli regression functions. The criterion of testing the hypothesis of two Bernoulli regression functions is constructed. The question as to its consistency is studied. The power asymptotics of the constructed criterion is also studied for certain types of close alternatives.

Assume that random variables $Y^{(i)}$, $i=1,2$, take two values: 1 and 0 with probabilities p_i ("success") and $1-p_i$ ("failure"), $i=1,2$, respectively. Assume that the probability of "success" p_i is a function of an independent variable $x\in[0,1]$, i.e., $p_i=p_i(x)=\mathbb{P}\{Y^{(i)}=1\mid x\}$ (see [2,3,8]). Let t_j, $j=1,\dots,n$, be points of a partition of the segment $[0,1]$:

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

Let $Y_i^{(1)}$ and $Y_i^{(2)}$, $i=1,\dots,n$, be mutually independent Bernoulli random variables with

$$\mathbb{P}\{Y_i^{(k)}=1\mid\ t_i\}=p_k(t_i)\ \text{ and }\ \mathbb{P}\{Y_i^{(k)}=0\mid\ t_i\}=1-p_k(t_i),\quad i=1,\dots,n,\ \ k=1,2.$$

Using the samples $Y_1^{(1)},\dots,Y_n^{(1)}$ and $Y_1^{(2)},\dots,Y_n^{(2)}$, it is required to test the hypothesis

$$H_0:\ p_1(x)=p_2(x)=p(x),\ \ x\in[0,1],$$

against a sequence of "close" alternatives:

$$H_{1n}: p_1(x)=p(x),\ \ p_2(x)=p(x)+\alpha_n u(x)+o(\alpha_n),$$

where α_n tends to 0 in a suitable way, $u(x)\neq 0$, $x\in[0,1]$, and the third term is $o(\alpha_n)$ uniformly with respect to $x\in[0,1]$.

The problem of comparison of two Bernoulli regression functions may appear in some applications, e.g., in the quantum bioanalyses carried out in pharmacology. In this case, x is a dose of medicine and $p(x)$ is the probability of efficiency of the dose x [3,6].

To test the hypothesis H_0 we use the statistic:

$$\begin{aligned}T_n&=\frac{1}{2}\,nb_n\int\limits_{\Omega_n(\tau)}\big[\widehat{p}_{1n}(x)-\widehat{p}_{2n}(x)\big]^2p_n^2(x)\,dx\\&=\frac{1}{2}\,nb_n\int\limits_{\Omega_n(\tau)}\big[p_{1n}(x)-p_{2n}(x)\big]^2\,dx,\\&\qquad\Omega_n(\tau)=[\tau b_n,1-\tau b_n],\ \ \tau>0,\end{aligned}$$

2020 *Mathematics Subject Classification.* 62G10, 62G20.
Key words and phrases. Bernoulli regression function; Power; Consistency; Limiting distribution.

where

$$\widehat{p}_{in}(x) = p_{in}(x) p_n^{-1}(x),$$
$$p_{in}(x) = (nb_n)^{-1} \sum_{j=1}^{n} K\Big(\frac{x-t_j}{b_n}\Big) Y_j^{(i)}, \quad i=1,2,$$
$$p_n(x) = (nb_n)^{-1} \sum_{j=1}^{n} K\Big(\frac{x-t_j}{b_n}\Big),$$

$K(x)$ is a distribution density, $b_n \to 0$ is a sequence of positive numbers, and $\widehat{p}_{in}(x)$ is a kernel estimate for the regression function [6, 9].

We assume that the kernel $K(x) \geq 0$ is chosen so that it is a function with bounded variation satisfying the following conditions: $K(x) = K(-x)$, $K(x) = 0$ for $|x| \geq \tau > 0$ and

$$\int K(x)\, dx = 1,$$

By $H(\tau)$, we denote the class of such functions.

We also introduce the following notation:

$$T_n^{(1)} = \frac{1}{2}\, nb_n \int\limits_{\Omega_n(\tau)} \big[\widetilde{p}_{1n}(x) - \widetilde{p}_{2n}(x)\big]^2\, dx,$$
$$\widetilde{p}_{in}(x) = p_{in}(x) - \boldsymbol{E} p_{in}(x), \quad i = 1, 2.$$

It is clear that

$$T_n^{(1)} = H_n + \frac{1}{2nb_n} \sum_{i=1}^{n} \varepsilon_i^2 Q_{ii}, \quad H_n = \frac{1}{nb_n} \sum_{1\leq i<j\leq n} \varepsilon_i \varepsilon_j Q_{ij},$$
$$\varepsilon_i = \varepsilon_{1i} - \varepsilon_{2i}, \quad \varepsilon_{ki} = Y_i^{(k)} - p_k(t_i), \quad k = 1, 2, \quad i = 1, \dots, n,$$
$$Q_{ij} = \psi_n(t_i, t_j), \quad \psi_n(u, v) = \int\limits_{\Omega_n(\tau)} K\Big(\frac{x-u}{b_n}\Big) K\Big(\frac{x-v}{b_n}\Big)\, dx.$$

It is easy to see that

$$\sigma_n^{-1}(T_n^{(1)} - \Delta_n) = \sum_{k=1}^{n} \xi_k^{(n)} + \frac{1}{2nb_n\sigma_n} \sum_{i=1}^{n} (\varepsilon_i^2 - \boldsymbol{E}\varepsilon_i^2) Q_{ii},$$
$$\Delta_n = \boldsymbol{E} T_n^{(1)}, \quad \sigma_n^2 = \boldsymbol{Var}\, H_n = (nb_n)^{-2} \sum_{k=2}^{n} d_k \sum_{i=1}^{k-1} d_i Q_{ik}^2,$$
$$d_i = d(t_i) = \boldsymbol{Var}\, \varepsilon_i, \quad i = 1, \dots, n,$$
$$\xi_k^{(n)} = \sum_{i=1}^{k-1} \eta_{ik}^{(n)}, \quad k = 2, \dots, n, \quad \xi_1^{(n)} = 0, \quad \xi_k^{(n)} = 0, \quad k > n,$$
$$\eta_{ij}^{(n)} = \frac{\varepsilon_i \varepsilon_j Q_{ij}}{nb_n\sigma_n}, \quad \mathcal{F}_k^{(n)} = \sigma(\varepsilon_1, \dots, \varepsilon_k),$$

i.e., $\mathcal{F}_k^{(n)}$ is a σ-algebra generated by random variables $\varepsilon_1, \dots, \varepsilon_k$ and $\mathcal{F}_0^{(n)} = (\varnothing, \Omega)$ in what follows, for simplicity, we use the notation $\xi_k^{(n)}$, $\eta_{ij}^{(n)}$ and $\mathcal{F}_k^{(n)}$ instead of ξ_k, η_{ij} $\mathcal{F}_k$.

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

Lemma 2 ([7])**.** *Let $K(x) \in H(\tau)$ and $p(x)$, $0 \leq x \leq 1$, be a function of bounded variation. If $nb_n \to \infty$, then*

$$\frac{1}{nb_n} \sum_{i=1}^{n} K^{\nu_1}\Big(\frac{x-t_i}{b_n}\Big) K^{\nu_2}\Big(\frac{y-t_i}{b_n}\Big) p^{\nu_3}(t_i)$$

$$= \frac{1}{b_n} \int_0^1 K^{\nu_1}\Big(\frac{x-u}{b_n}\Big) K^{\nu_2}\Big(\frac{y-u}{b_n}\Big) p^{\nu_3}(u)\, du + O\Big(\frac{1}{nb_n}\Big),$$

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

Lemma 3. *Let* $K(x) \in H(\tau)$, $p(x) \in C^1[0,1]$ *and let* $u(x)$ *be a continuous function on* $[0,1]$. *If* $nb_n^2 \to \infty$ *and* $\alpha_n = n^{-1/2} b_n^{-1/4}$, *then, for the hypothesis* H_{1n}

$$b_n^{-1}\sigma_n^2 \longrightarrow \sigma^2(p) = 2\int_0^1 p^2(x)(1-p(x))^2\, dx \int_{|x|\le 2\tau} K_2^2(x)\, dx \tag{1}$$

and

$$b_n^{-1/2}(\Delta_n - \Delta(p)) = O(b_n^{1/2}) + O(\alpha_n b_n^{-1/2}) + O\Big(\frac{1}{nb_n^{3/2}}\Big), \tag{2}$$

where

$$\Delta_n = \boldsymbol{E} T_n^{(1)}, \quad \Delta(p) = \int_0^1 p(x)(1-p(x))\, dx \int_{|x|\le\tau} K^2(u)\, du, \quad K_2 = K * K,$$

and $*$ *denotes the operation of convolution.*

The following statement is true:

Theorem 1. *Let* $K(x) \in H(\tau)$ *and* $p(x), u(x) \in C^1[0,1]$. *If* $nb_n^2 \to \infty$ *and* $\alpha_n = n^{-1/2} b_n^{-1/4}$, *then, for the hypothesis* H_{1n},

$$b_n^{-1/2}(T_n - \Delta(p))\sigma^{-1}(p) \xrightarrow{d} N(a,1),$$

where $\Delta(p)$ *and* $\sigma^2(p)$ *are defined in Lemma* 3, $\xrightarrow{d}$ *denotes the convergence in distribution,* $N(a,1)$ *is a random variable having normal distribution with parameters* $(a,1)$, *and*

$$a = \frac{1}{2\sigma(p)} \int_0^1 u^2(x)\, dx.$$

Proof. We have

$$T_n = T_n^{(1)} + L_n^{(1)} + L_n^{(2)},$$

where

$$L_n^{(1)} = nb_n \int_{\Omega_n(\tau)} \big[\widetilde{p}_{1n}(x) - \widetilde{p}_{2n}(x)\big]\big[\boldsymbol{E}p_{1n}(x) - \boldsymbol{E}p_{2n}(x)\big]\, dx,$$

$$L_n^{(2)} = \frac{1}{2}\, nb_n \int_{\Omega_n(\tau)} \big[\boldsymbol{E}p_{1n}(x) - \boldsymbol{E}p_{2n}(x)\big]^2\, dx.$$

By virtue of Lemma 2, we conclude that

$$b_n^{-1/2} L_n^{(2)} = \frac{1}{2}\, nb_n^{1/2}\alpha_n^2 \int_{\Omega_n(\tau)} \Big\{\frac{1}{b_n}\int_0^1 K\Big(\frac{x-t}{b_n}\Big) u(t)\, dt + O\Big(\frac{1}{nb_n}\Big)\Big\}^2 dx. \tag{3}$$

Since $\big[\frac{x-1}{b_n}, \frac{x}{b_n}\big] \supset [-\tau,\tau]$ for all $x \in \Omega_n(\tau)$, it follows from (3) that

$$b_n^{-1/2} L_n^{(2)} = \frac{1}{2}\, nb_n^{1/2}\alpha_n^2 \int_{\Omega_n(\tau)} \Big[\int_{-\tau}^{\tau} K(t) u(x - b_n t)\, dt + O\Big(\frac{1}{nb_n}\Big)\Big]^2 dx. \tag{4}$$

Further, since $u(x)\in C^1[0,1]$, in view of (4), we get

$$b_n^{-1/2}L_n^{(2)}\longrightarrow \frac{1}{2}\int\limits_0^1 u^2(t)\,dt. \tag{5}$$

We now show that

$$b_n^{-1/2}L_n^{(1)}\xrightarrow{\boldsymbol{P}}0.$$

Thus, we have

$$b_n^{-1/2}L_n^{(1)}=\frac{1}{2}\,nb_n^{1/2}\int\limits_{\Omega_n(\tau)}\widetilde{p}_{1n}(x)\big(\boldsymbol{E}p_{1n}(x)-\boldsymbol{E}p_{2n}(x)\big)\,dx$$

$$-\frac{nb_n^{1/2}}{2}\int\limits_{\Omega_n(\tau)}\widetilde{p}_{2n}(x)\big(\boldsymbol{E}p_{1n}(x)-\boldsymbol{E}p_{2n}(x)\big)\,dx=I_n^{(1)}+I_n^{(2)}. \tag{6}$$

It is clear that

$$\begin{aligned}\boldsymbol{E}|I_n^{(1)}|&\le\big(\boldsymbol{E}(I_n^{(1)})^2\big)^{1/2}\\&=\frac{1}{2}\,nb_n^{1/2}\bigg[\boldsymbol{E}\bigg(\int\limits_{\Omega_n(\tau)}\widetilde{p}_{1n}(x)\big(\boldsymbol{E}p_{1n}(x)-\boldsymbol{E}p_{2n}(x)\big)\,dx\bigg)^2\bigg]^{1/2}\\&=\frac{1}{2}\,nb_n^{1/2}\bigg[\int\limits_{\overline{\Omega}_n(\tau)}\operatorname{cov}\big(p_{1n}(x_1),p_{1n}(x_2)\big)\big(\boldsymbol{E}p_{1n}(x_1)-\boldsymbol{E}p_{2n}(x_1)\big)\\&\qquad\times\big(\boldsymbol{E}p_{1n}(x_2)-\boldsymbol{E}p_{2n}(x_2)\big)\,dx_1\,dx_2\bigg]^{1/2},\ \ \overline{\Omega}_n(\tau)=\Omega_n(\tau)\times\Omega_n(\tau).\end{aligned}$$

It is easy to see that

$$\begin{aligned}&\operatorname{cov}\big(p_{1n}(x_1),p_{1n}(x_2)\big)\\&=\frac{1}{(nb_n)^2}\sum_{i=1}^n K\Big(\frac{x_1-t_i}{b_n}\Big)K\Big(\frac{x_2-t_i}{b_n}\Big)p_1(t_i)(1-p_1(t_i)).\end{aligned}$$

By virtue of Lemma 2, we find

$$\begin{aligned}&\operatorname{cov}\big(p_{1n}(x_1),p_{1n}(x_2)\big)\\&=n^{-1}b_n^{-2}\int\limits_0^1 K\Big(\frac{x_1-u}{b_n}\Big)K\Big(\frac{x_2-u}{b_n}\Big)p_1(u)(1-p_1(u))\,du+O\Big(\frac{1}{(nb_n)^2}\Big).\end{aligned}$$

Hence,

$$\begin{aligned}\boldsymbol{E}|I_n^{(1)}|&\le\frac{1}{2}\,nb_n^{1/2}\bigg\{\int\limits_{\overline{\Omega}_n(\tau)}\bigg[\frac{1}{nb_n^2}\\&\times\int\limits_0^1 K\Big(\frac{x_1-u}{b_n}\Big)K\Big(\frac{x_2-u}{b_n}\Big)p_1(u)(1-p_1(u))\,du+\frac{1}{(nb_n)^2}\bigg]\\&\times\big(\boldsymbol{E}p_{1n}(x_1)-\boldsymbol{E}p_{2n}(x_1)\big)\big(\boldsymbol{E}p_{1n}(x_2)-\boldsymbol{E}p_{2n}(x_2)\big)\,dx_1\,dx_2\bigg\}^{1/2}\\&\le c_3\sqrt{n}\,b_n^{1/2}\alpha_n=c_3\,\frac{1}{\sqrt{n}\,\alpha_n}\longrightarrow 0,\end{aligned}$$

because

$$\sqrt{n}\,\alpha_n = \frac{1}{b_n^{1/4}} \longrightarrow \infty.$$

Therefore, $I_n^{(1)} \xrightarrow{\boldsymbol{P}} 0$. Similarly, we prove that $I_n^{(2)} \xrightarrow{\boldsymbol{P}} 0$.

By using (6), we get

$$b_n^{-1/2} L_n^{(1)} \xrightarrow{\boldsymbol{P}} 0. \tag{7}$$

To prove the theorem, it remains to show that

$$\frac{T_n^{(1)} - \Delta_n}{\sigma_n} \xrightarrow{d} N(0,1).$$

We have

$$\frac{T_n^{(1)} - \Delta_n}{\sigma_n} = K_n^{(1)} + K_n^{(2)},$$

where

$$K_n^{(1)} = \sum_{k=1}^n \xi_k, \quad K_n^{(2)} = \frac{\sum\limits_{i=1}^n (\varepsilon_i^2 - \boldsymbol{E}\varepsilon_i^2) Q_{ii}}{2nb_n\sigma_n}\,.$$

We now show that $K_n^{(2)} \xrightarrow{\boldsymbol{P}} 0$. Indeed,

$$\boldsymbol{Var}(K_n^{(2)}) = (2nb_n\sigma_n)^{-2} \sum_{i=1}^n \boldsymbol{Var}\varepsilon_i^2 Q_{ii}^2$$
$$= (2nb_n\sigma_n)^{-2} \sum_{i=1}^n \bigg(\sum_{k=1}^2 p_k(t_i)(1-p_k(t_i))\big[1 - 3p_k(t_i)(1-p_k(t_i))\big] \bigg) Q_{ii}^2.$$

Since $Q_{ii} \le c_4 b_n$ $b_n^{-1}\sigma_n^2 \longrightarrow \sigma^2(p)$ as $n \to \infty$, this yields

$$\boldsymbol{Var}(K_n^{(2)}) \le c_5\,\frac{1}{nb_n}\,.$$

Thus, $K_n^{(2)} \xrightarrow{\boldsymbol{P}} 0$.

We now prove that $K_n^{(1)} \xrightarrow{d} N(0,1)$. To this end, we show that it is possible to apply Corollaries 2 and 6 of Theorem 2 in [4]. It is necessary to check the validity of conditions imposed in these statements and guaranteeing the asymptotic normality of a square-integrable martingale difference and to take into account the fact that, according to Lemma 1, the sequence $\{\xi_k, \mathcal{F}_k\}_{k\ge1}$ is, in fact, a square-integrable martingale difference.

It is easy to see that $\sum\limits_{k=1}^n \boldsymbol{E}\xi_k^2 = 1$. The asymptotic normality of $K_n^{(1)}$ is realized whenever

$$\sum_{k=1}^n \boldsymbol{E}\Big[\xi_k^2 I\big(|\xi_k| \ge \varepsilon\big) \mid \mathcal{F}_{k-1}\Big] \xrightarrow{\boldsymbol{P}} 0 \tag{8}$$

and

$$\sum_{k=1}^n \xi_k^2 \xrightarrow{\boldsymbol{P}} 1 \tag{9}$$

as $n \to \infty$. In [4], it is proved that, under the conditions

$$\sup_{1\le k\le n} |\xi_k| \xrightarrow{\boldsymbol{P}} 0$$

and (9), condition (8) is also satisfied.

Note that, for $\varepsilon > 0$, we have

$$P\bigg\{ \sup_{1\le k\le n} |\xi_k| \ge \varepsilon \bigg\} \le \varepsilon^{-4} \sum_{k=1}^n \boldsymbol{E}\xi_k^4.$$

Hence, by virtue of relation (11) presented in what follows, in order to prove

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

it remains to check the validity of condition (9). To this end, it suffices to show that

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

i.e., since $\sum\limits_{k=1}^n \boldsymbol{E}\xi_k^2 = 1$, we get

$$\boldsymbol{E}\Big(\sum_{k=1}^n \xi_k^2\Big)^2 = \sum_{k=1}^n \boldsymbol{E}\xi_k^4 + 2\sum_{1\le k_1<k_2\le n} \boldsymbol{E}\xi_{k_1}^2\xi_{k_2}^2 \longrightarrow 1. \tag{10}$$

We now prove (10). Taking into account the definitions of η_{ik} and ξ_k, we obtain

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

where

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

$$I_n^{(2)} = \frac{3}{(nb_n)^4\sigma_n^4}\sum_{k=2}^n \sum_{i\ne j} \boldsymbol{E}\varepsilon_j^2 \boldsymbol{E}\varepsilon_i^2 Q_{jk}^2 Q_{ik}^2.$$

Since

$$Q_{ij} \le c_6 b_n, \quad \boldsymbol{E}\varepsilon_j^4 \le 8\sum_{k=1}^2 p_k(t_j)\big(1-p_k(t_j)\big)\Big[1 - 3p_k(t_j)\big(1-p_k(t_j)\big)\Big] \le 4,$$

$$\boldsymbol{E}\varepsilon_j^2 \le \frac{1}{2}\,, \quad |\boldsymbol{E}\varepsilon_j^3| \le \sum_{k=1}^2 p_k(t_j)\big(1-p_k(t_j)\big)\Big[\big(1-p_k(t_j)\big)^2 + p_k^2(t_j)\Big] \le 1$$

and $b_n^{-1}\sigma_n^2 \longrightarrow \sigma^2(p)$, we find

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

Hence,

$$\sum_{k=1}^n \boldsymbol{E}\xi_k^4 \longrightarrow 0 \ \text{ for } \ n \to \infty. \tag{11}$$

Further, it follows from the definition of ξ_i 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

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

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

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

Therefore,

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

where

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

We now consider $A_n^{(3)}$. By using the definition of η_{ij}, we can easily show that $\boldsymbol{E} B_{k_1k_2}^{(3)} = 0$ and, hence,

$$A_n^{(3)} = 0. \tag{12}$$

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

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

Since $\boldsymbol{E}|\varepsilon_i^3| \le 1$ and $Q_{ij} \le c_6 b_n$, we get

$$|\boldsymbol{E} B_{k_1k_2}^{(2)}| \le c_6 \, \frac{k_1-1}{(n\sigma_n)^4}\,.$$

Further, since

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

we obtain

$$|A_n^{(2)}| \le c_7 \, \frac{n^3}{n^4\sigma_n^4} = c_7 \, \frac{1}{nb_n^2(b_n^{-1}\sigma_n^2)^2} = O\Big(\frac{1}{nb_n^2}\Big). \tag{13}$$

We now establish that $A_n^{(1)} \to 1$ as $n \to \infty$. It is clear that

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

where

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

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

It follows from the definition of σ_n^2 that

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

Furthermore,

$$\sum_{k=2}^{n} \Big(\sum_{i=1}^{k-1} \boldsymbol{E}\eta_{ik}^2 \Big)^2 \le c_8 \, \frac{b_n^4 n^3}{(nb_n)^4\sigma_n^4} = O\Big(\frac{1}{nb_n^2}\Big).$$

This yields

$$S_n^{(1)} = 1 + O\Big(\frac{1}{nb^2}\Big). \tag{14}$$

Further, we show that $S_n^{(2)} \to 0$. The quantity $S_n^{(2)}$ can be rewritten in the form

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

It is easy to see that

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

However, since

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

we conclude that

$$S_n^{(2)} = O\Big(\frac{1}{n\sigma_n^4}\Big) = O\Big(\frac{1}{nb_n^2}\Big). \tag{15}$$

Hence, according to (14) and (15), we find

$$A_n^{(1)} = 1 + O\Big(\frac{1}{nb_n^2}\Big). \tag{16}$$

Finally, we show that $A_n^{(4)} \to 0$ as $n \to \infty$. By using the definition of η_{ij} and the inequalities $Q_{ij} \geq 0$ and

$$\boldsymbol{E}\varepsilon_i^2 = d(t_i) \leq \frac{1}{2}\,,$$

we obtain

$$\boldsymbol{E}B_{k_1k_2}^{(4)} = 4 \sum_{1\leq t<s\leq k_1-1} \boldsymbol{E}\eta_{sk_1}\eta_{tk_1}\eta_{sk_2}\eta_{tk_2}$$

$$\leq \frac{c_8}{n^4 b_n^4 \sigma_n^4} \sum_{1\leq t<s\leq k_1-1} Q_{sk_1}Q_{tk_1}Q_{sk_2}Q_{tk_2}.$$

Thus,

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

where

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

At the same time,

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

Therefore,

$$A_n^{(4)} \leq c_{10}\,\frac{1}{n^2 b_n^4 \sigma_n^4} \sum_{k_1,k_2=1}^{n} \Big[\int\limits_{\Omega(\tau)}\int\limits_{\Omega_n(\tau)} K\Big(\frac{x-x_{k_1}}{b_n}\Big)K\Big(\frac{y-x_{k_2}}{b_n}\Big)$$

$$\times\frac{1}{n}\sum_{i=1}^{n} K\Big(\frac{x-x_i}{b_n}\Big)K\Big(\frac{y-x_i}{b_n}\Big)\,dx\,dy\Big)\Big]^2. \tag{17}$$

Further, in view of Lemma 2, it follows from (17) that

$$A_n^{(4)} \leq \frac{c_{11}}{b_n^4\sigma_n^4} \sum_{k_1,k_2=1}^{n} \Big\{\frac{1}{n}\int\limits_0^1 \int\limits_{\Omega_n(\tau)}\int\limits_{\Omega_n(\tau)} K\Big(\frac{x-x_{k_1}}{b_n}\Big)K\Big(\frac{y-x_{k_2}}{b_n}\Big)$$

$$\times K\Big(\frac{x-u}{b_n}\Big)K\Big(\frac{y-u}{b_n}\Big)\,du\,dx\,dy\Big\}^2 + O\Big(\frac{1}{nb_n^2}\Big). \tag{18}$$

In relation (18), we now apply Lemma 2 once again. This yields

$$A_n^{(4)} \leq \frac{c_{12}}{b_n^4\sigma_n^4} \int\limits_0^1\int\limits_0^1\int\limits_0^1\int\limits_0^1 \psi_n(u_1,v_2)\psi_n(u_1,v_1)$$

$$\times\psi_n(u_2,v_1)\psi_n(u_2,v_2)\,du_1\,du_2\,dv_1\,dv_2, \tag{19}$$

where

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

We now estimate the integral in (19). Since $\left[\frac{x-1}{b_n}, \frac{x}{b_n}\right] \supseteq [-\tau, \tau]$ for all $x \in \Omega_n(\tau)$, we get

$$\begin{aligned}
&\int_0^1 \psi_n(u_1, v_2)\psi_n(u_1, v_1)\, du_1 \\
&= b_n \int_{\overline{\Omega}_n(\tau)} K\Big(\frac{t - v_2}{b_n}\Big) K\Big(\frac{z - v_1}{b_n}\Big) K_2\Big(\frac{z - t}{b_n}\Big)\, dt\, dz \\
&\le c_{13} b_n^3, \quad K_2 = K * K, \ \ \overline{\Omega}_n(\tau) = \Omega_n(\tau) \times \Omega_n(\tau).
\end{aligned}$$

Hence,

$$A_n^{(4)} \le c_{14} \frac{1}{b_n \sigma_n^4} \int_0^1 \int_0^1 \int_0^1 \psi_n(u_2, v_1)\psi_n(u_2, v_2)\, du_2\, dv_1\, dv_2 + O\Big(\frac{1}{n b_n^2}\Big). \tag{20}$$

Further, in a similar way, we derive the following result from (20):

$$\begin{aligned}
A_n^{(4)} &\le c_{15} \frac{b_n^4}{b_n \sigma_n^4} + O\Big(\frac{1}{n b_n^2}\Big) = O\Big(\frac{b_n^4}{b_n^3 (b_n^{-1} \sigma_n^2)^2}\Big) + O\Big(\frac{1}{n b_n^2}\Big) \\
&= O(b_n) + O\Big(\frac{1}{n b_n^2}\Big).
\end{aligned} \tag{21}$$

Combining relations(12), (13), (16) and (21), we conclude that

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

In view of relation (11), this yields that

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

Hence,

$$\frac{T_n^{(1)} - \Delta_n}{\sigma_n} \xrightarrow{d} N(0, 1). \tag{22}$$

Further, by using the representation $T_n = T_n^{(1)} + L_n^{(1)} + L_n^{(2)}$, Lemma 3 and relations (5), (7), and (22) we get

$$b_n^{-1/2} \Big(\frac{T_n - \Delta(p)}{\sigma(p)}\Big) \xrightarrow{d} N\Big(\frac{1}{2\sigma(p)} \int_0^1 u^2(x)\, dx, 1\Big).$$

Theorem 1 is proved. □

Corollary 1. *Let $K(u) \in H(\tau)$ and $p(x) \in C^1[0,1]$. If $n b_n^2 \to \infty$, then the following relation is true for the hypothesis H_0:*

$$b_n^{-1/2}(T_n - \Delta(p))\sigma^{-1}(p) \xrightarrow{d} N(0, 1). \tag{23}$$

As an important application of Corollary 1, we construct a criterion for the testing of a simple hypothesis H_0 of equality of two Bernoulli regression functions $p_1(x) = p_2(x) = p(x)$, where the function $p(x)$ is completely defined. The critical domain is determined by the inequality

$$T_n \ge d_n(\alpha) = \Delta(p) + b_n^{1/2} \sigma(p) \lambda_\alpha,$$

where $\Phi(\lambda_\alpha) = 1 - \alpha$ and $\Phi(\lambda)$ is the standard normal distribution.

Corollary 2. *Let $K(u) \in H(\tau)$ and $p(x), u(x) \in C^1[0,1]$. If $n b_n^2 \to \infty$ and $\alpha_n = n^{-1/2} b_n^{-1/4}$, then the local behavior of the power $\boldsymbol{P}_{H_{1n}}(T_n \ge d_n(\alpha))$ has the form*

$$\boldsymbol{P}_{H_{1n}}(T_n \ge d_n(\alpha)) \longrightarrow 1 - \Phi\Big(\lambda_\alpha - \frac{A(u)}{\sigma(p)}\Big),$$

where

$$A(u)=\frac{1}{2}\int\limits_0^1 u^2(x)\,dx>0.$$

We now assume that $p(x)$ is not defined by the hypothesis (i.e., we testing a ***composite*** hypothesis).

In this case, it is impossible to apply inequality (1) directly. First, it is necessary to replace the unknown parameters $\Delta(p)$ and $\sigma^2(p)$ appearing in (23) by certain estimates $\widetilde{\Delta}_n$ and $\widetilde{\sigma}_n^2$, respectively. As the estimates $\Delta(p)$ and $\sigma^2(p)$, we take the following statistics:

$$\widetilde{\Delta}_n=\int\limits_{\Omega_n(\tau)}\lambda_n(x)\,dx\int\limits_{|x|\le\tau}K^2(x)\,dx,$$

$$\widetilde{\sigma}_n^2=2\int\limits_{\Omega_n(\tau)}\lambda_n^2(x)\,dx\int\limits_{|x|\le 2\tau}K_2^2(x)\,dx,$$

$$\lambda_n(x)=\frac{1}{2}\Big[p_{1n}(x)\big(p_n(x)-p_{1n}(x)\big)+p_{2n}(x)\big(p_n(x)-p_{2n}(x)\big)\Big].$$

We now show that

$$b_n^{-1/2}(\widetilde{\Delta}_n-\Delta(p))\xrightarrow{\boldsymbol{P}}0,\quad \widetilde{\sigma}_n^2\xrightarrow{\boldsymbol{P}}\sigma^2(p). \tag{24}$$

Indeed, since

$$p_n(x)=1+O\Big(\frac{1}{nb_n}\Big)$$

uniformly with respect to $x\in\Omega_n(\tau)$ and $|p_{in}(x)|\le c_{16}$, $x\in[0,1]$, $i=1,2$, we find

$$\begin{aligned}
&b_n^{-1/2}\boldsymbol{E}|\widetilde{\Delta}_n-\Delta(p)|\\
&\le c_{17}b_n^{-1/2}\Big[\int\limits_{\Omega_n(\tau)}\Big(\boldsymbol{E}\big(p_{1n}(x)-\boldsymbol{E}p_{1n}(x)\big)^2\Big)^{1/2}dx\\
&+\int\limits_{\Omega_n(\tau)}\Big(\boldsymbol{E}\big(p_{2n}(x)-\boldsymbol{E}p_{2n}(x)\big)^2\Big)^{1/2}dx\Big]\\
&+b_n^{-1/2}\int\limits_{\Omega_n(\tau)}\big|\boldsymbol{E}p_{1n}(x)-p(x)\big|\,dx+b_n^{-1/2}\int\limits_{\Omega_n(\tau)}\big|\boldsymbol{E}p_{2n}(x)-p(x)\big|\,dx.
\end{aligned}$$

Further, by using Lemma 2 and taking into account the facts that $p(x)\in C^1[0,1]$ and $\big[\frac{x-1}{b_n},\frac{x}{b_n}\big]\supset[-\tau,\tau]$ for all $x\in\Omega_n(\tau)$, we immediately conclude that

$$\begin{aligned}
&b_n^{-1/2}\boldsymbol{E}|\widetilde{\Delta}_n-\Delta(p)|\\
&\le c_{18}b_n^{-1/2}\Big\{\int\limits_{\Omega_n(\tau)}\Big[\frac{1}{nb_n}\,\frac{1}{b_n}\int\limits_0^1K^2\Big(\frac{x-u}{b_n}\Big)p(u)(1-p(u))\,du+O\Big(\frac{1}{(nb_n)^2}\Big)\Big]^{1/2}\\
&+O(b_n)+O\Big(\frac{1}{nb_n}\Big)\Big\}=O\Big(\frac{1}{\sqrt{n}\,b_n}\Big)+O(b_n^{1/2})+O\Big(\frac{1}{nb^{3/2}}\Big).
\end{aligned}$$

Hence, $b_n^{-1/2}(\widetilde{\Delta}_n-\Delta(p))\xrightarrow{\boldsymbol{P}}0$. Similarly, we can show that $\widetilde{\sigma}_n^2\xrightarrow{\boldsymbol{P}}\sigma^2(p)$.

Theorem 2. *Let $K(x)\in H(\tau)$ and $p_1(x)=p_2(x)=p(x)\in C^1[0,1]$. If $nb_n^2\to\infty$, then, as $n\to\infty$,*

$$b_n^{-1/2}(T_n-\widetilde{\Delta}_n)\widetilde{\sigma}_n^{-1}\xrightarrow{d}N(0,1).$$

The *proof* follows from (23) and (24).

Theorem 2 enables us to construct an asymptotic criterion for the testing of the ***composite*** hypothesis

$$H_0:\ \ p_1(x)=p_2(x),\ \ x\in[0,1].$$

The critical domain for the testing of this hypothesis is given by the inequality

$$T_n \geq \widetilde{d}_n(\alpha) = \widetilde{\Delta}_n + b_n^{-1/2}\widetilde{\sigma}_n\lambda_\alpha, \quad \Phi(\lambda_\alpha) = 1 - \alpha. \tag{25}$$

Now let us investigate the asymptotic property of criterion (25) (i.e., the behavior of the power function as $n \to \infty$).

Theorem 3. *Let $K(x) \in H(\tau)$, $p_1(x), p_2(x) \in C^1[0,1]$. If $nb_n^2 \to \infty$, then*

$$\gamma_n(p_1, p_2) = \boldsymbol{P}_{H_1}\big(T_n \geq \widetilde{d}_n(\alpha)\big) \longrightarrow 1$$

as $n \to \infty$. Any pair $(p_1(x), p_2(x))$, $0 \leq p_i(x) \leq 1$, $p_i(x) \in C^1[0,1]$, $i = 1, 2$, such that $p_1(x) \neq p_2(x)$ at at least one point x, $x \in [0,1]$. is an alternative of the hypothesis H_1.

Proof. Denote

$$\overline{T}_n = \frac{1}{2}\, nb_n \int\limits_{\Omega_n} \big(\overline{p}_{1n}(x) - \overline{p}_{2n}(x)\big)^2\, dx,$$

$$\overline{p}_{in}(x) = p_{in}(x) - \boldsymbol{E}p_{in}(x), \quad i = 1, 2.$$

By analogy with (1), (2) and (24), we can readily show that the following is true for the hypothesis H_1

$$\begin{gathered}
b_n^{-1}\sigma_n^2 \longrightarrow \sigma^2(p_1, p_2) = 2\int\limits_0^1 d^2(x)\, dx \int\limits_{|x| \leq 2\tau} K_2^2(x)\, dx, \\
\widetilde{\sigma}_n^2 \xrightarrow{\boldsymbol{P}} \sigma^2(p_1, p_2), \quad \widetilde{\Delta}_n \xrightarrow{\boldsymbol{P}} \Delta(p_1, p_2), \quad \boldsymbol{E}\overline{T}_n \longrightarrow \Delta(p_1, p_2), \\
\Delta(p_1, p_2) = \int\limits_0^1 d(x)\, dx \int\limits_{|x| \leq \tau} K^2(x)\, dx, \\
d(x) = \frac{1}{2}\sum_{k=1}^{2} p_k(x)(1 - p_k(x)).
\end{gathered} \tag{26}$$

Further, in view of Lemma 2 and the fact that $\left[\frac{x-1}{b_n}, \frac{x}{b_n}\right] \supset [-\tau, \tau]$, $x \in \Omega_n(\tau)$ we obtain

$$\int\limits_{\Omega_n} \big(\boldsymbol{E}p_{1n}(x) - \boldsymbol{E}p_{2n}(x)\big)^2\, dx$$

$$= \int\limits_{\Omega_n} \bigg(\int\limits_{-\tau}^{\tau} K(t)\big(p_1(x - b_n(u)) - p_2(x - b_n(u))\big)^2\, du\bigg)\, dx + O\Big(\frac{1}{nb_n}\Big).$$

According to the condition $p_1(x), p_2(x) \in C^1[0,1]$, we get

$$\int\limits_{\Omega_n} \big(\boldsymbol{E}p_{1n}(x) - \boldsymbol{E}p_{2n}(x)\big)^2\, dx = \int\limits_0^1 \big(p_1(x) - p_2(x)\big)^2\, dx + O(b_n) + O\Big(\frac{1}{nb_n}\Big). \tag{27}$$

By using (26) and (27), after simple transformations, we find

$$\begin{gathered}
\gamma_n(p_1, p_2) = \boldsymbol{P}_{H_1}\bigg[\frac{\overline{T}_n - \boldsymbol{E}\overline{T}_n}{\sigma_n} \\
\geq -nb_n^{1/2}\bigg(\int\limits_0^1 \big(p_1(x) - p_2(x)\big)^2\, dx + o_p(1)\bigg)\bigg].
\end{gathered} \tag{28}$$

Finally, since

$$(\overline{T}_n - \boldsymbol{E}\overline{T}_n)\sigma_n^{-1} \xrightarrow{d} N(0,1)$$

(the proof of this statement is similar to the proof of (22)) and $nb_n^{1/2} \to \infty$, it follows from (28) that $\gamma_n(p_1, p_2) \to 1$ as $n \to \infty$. □

References

1. M. Aerts, N. Veraverbeke, Bootstrapping a nonparametric polytomous regression model. *Math. Methods Statist.* **4** (1995), no. 2, 189–200.
2. J. B. Copas, Plotting p against x. *Appl. Statist.* **32** (1983), no. 2, 25–31.
3. S. Efromovich, *Nonparametric Curve Estimation.* Methods, theory, and applications. Springer Series in Statistics. Springer-Verlag, New York, 1999.
4. R. Sh. Lipcer, A. N. Shirjaev, A functional central limit theorem for semimartingales. (Russian) *Teor. Veroyatnost. i Primenen.* **25** (1980), no. 4, 683–703.
5. H.-G. Müller, T. Schmitt, Kernel and probit estimates in quantal bioassay. *J. Amer. Statist. Assoc.* **83** (1988), no. 403, 750–759.
6. E. A. Nadaraya, On a regression estimate. (Russian) *Teor. Verojatnost. i Primenen.* **9** (1964), 157–159.
7. E. Nadaraya, P. Babilua, G. Sokhadze, Estimation of a distribution function by an indirect sample. *Ukr. Math. J.* **62** (2010), no. 12, 1642–1658.
8. 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.
9. G. S. Watson, Smooth regression analysis. *Sankhya Ser. A* **26** (1964), 359–372.

(Received 28.10.2019)

Department of Mathematics, Faculty of Exact and Natural Sciences, I. Javakhishvili Tbilisi State University, 3 University Str., Tbilisi 0143, Georgia

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

E-mail address: elizbar.nadaraya@tsu.ge