Georgian Mathematical Journal
1(1994), No. 2, 197-212

LIMIT DISTRIBUTION OF THE INTEGRATED SQUARED ERROR OF TRIGONOMETRIC SERIES REGRESSION ESTIMATOR

E. NADARAYA

Abstract. Limit distribution is studied for the integrated squared error of the projection regression estimator (2) constructed on the basis of independent observations (1). By means of the obtained limit theorems, a test is given for verifying the hypothesis on the regression, and the power of this test is calculated in the case of Pitman alternatives.

Let observations $Y_1, Y_2, \dots, Y_n$ be represented as

$$Y_i = \mu(x_i) + \varepsilon_i, \quad i = \overline{1,n}, \tag{1}$$

where $\mu(x)$, $x \in [-\pi, \pi]$, is the unknown regression function to be estimated by observations Y_i; x_i, $i = \overline{1,n}$, are the known numbers, and $-\pi = x_0 < x_1 < \cdots < x_n \le \pi$, ε_i, $i = \overline{1,n}$, are independent equidistributed random variables; $E\varepsilon_1 = 0$, $E\varepsilon_1^2 = \sigma^2$, and $E\varepsilon_1^4 < \infty$.

The problem of nonparametric estimation of the regression function $\mu(x)$ for the model (1) has a recent history and has been treated only in few papers. In particular, a kernel estimator of the Rosenblatt–Parzen type for $\mu(x)$ was proposed for the first time in [1].

Assume that $\mu(x)$ is representable as a converging series in $L_2(-\pi, \pi)$ with respect to the orthonormal trigonometric system

$$\left\{(2\pi)^{-1/2},\ \pi^{-1/2}\cos ix,\ \pi^{-1/2}\sin ix\right\}_{i=1}^{\infty}.$$

Consider the estimator of the function $\mu(x)$ constructed by the projection method of N.N. Chentsov [2]

$$\mu_{nN}(x) = \frac{a_{0n}}{2} + \sum_{i=1}^{N} a_{in}\cos ix + b_{in}\sin ix, \tag{2}$$

1991 *Mathematics Subject Classification.* 62G07.

where $N = N(n) \to \infty$ for $n \to \infty$ and

$$a_{in} = \frac{1}{\pi} \sum_{j=1}^{n} Y_j \Delta_j \cos i x_j, \quad b_{in} = \frac{1}{\pi} \sum_{j=1}^{n} Y_j \Delta_j \sin i x_j,$$
$$\Delta_j = x_j - x_{j-1}, \quad j = \overline{1,n}, \quad i = \overline{0,N}.$$

The estimator (2) can be rewritten in a more compact way as

$$\mu_{nN} = \sum_{j=1}^{n} Y_j \Delta_j K_N(x - x_j),$$

where $K_N(u) = \frac{1}{2\pi} \sum\limits_{|r| \le N} e^{iru}$ is the Dirichlet kernel.

In [3], p.347, N.V. Smirnov considered estimators of the type (2) for a specially chosen class of functions $\mu(x)$ in the case of equidistant points $x_j \in [-\pi, \pi]$ and of independent and normally distributed observation errors ε_i. In [4] an estimator of the type (2) is obtained, which is asymptotically equivalent to projection estimators which are optimal in the sense of some accuracy criterion. The asymptotics of the mean value of the integrated squared error of the estimator (2) is considered in [5].

It is of interest to investigate the limit distribution of the integrated squared error

$$\int_{-\pi}^{\pi} [\mu_{nN}(x) - \mu(x)]^2 dx,$$

which is the goal pursued in this paper. The method used to prove the theorems below is based on the functional limit theorem for a sequence of semimartingales [6].

Denote

$$U_{nN} = \frac{n}{2\pi(2N+1)} \int_{-\pi}^{\pi} \left[\mu_{nN}(x) - E\mu_{nN}(x)\right]^2 dx,$$

$$Q_{ir} = \Delta_i \Delta_r K_N(x_i - x_r), \quad \sigma_{nN}^2 = \frac{n^2 \sigma^4}{\pi^2 (2N+1)^2} \sum_{r=2}^{n} \sum_{j=1}^{r-1} Q_{jr}^2,$$

$$\eta_{ik} = \frac{n}{\pi(2N+1)\sigma_{nN}} \varepsilon_i \varepsilon_k Q_{ik},$$

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

and assume that $\mathcal{F}_k$ is σ-algebra generated by random variables $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k$, $\mathcal{F}_0 = (\phi, \Omega)$.

Lemma 1 ([7], p.179). *The stochastic sequence $(\xi_k, \mathcal{F}_k)_{k \ge 1}$ is a martingale-difference.*

Lemma 2. *Let $p(x)$ be the known positive continuously differentiable distribution density on $[-\pi,\pi]$, and points x_i be chosen from the relation $\int_{-\pi}^{x_i} p(u)\,du = \frac{i}{n}$, $i=\overline{1,n}$.*

If $\frac{N\ln N}{n} \to 0$ for $n\to\infty$, then

$$EU_{nN} = \theta_1 + O\left(\frac{N\ln N}{n}\right), \quad \theta_1 = \frac{\sigma^2}{(2\pi)^2}\int_{-\pi}^{\pi} p^{-1}(u)\,du, \tag{3}$$

$$(2N+1)\sigma_{nN}^2 \to \theta_2 = \frac{\sigma^4}{4\pi^3}\int_{-\pi}^{\pi} p^{-2}(u)\,du. \tag{4}$$

Proof. From the definition of x_i we easily obtain

$$\Delta_i = \frac{1}{np(x_i)}\left[1+O\Big(\frac{1}{n}\Big)\right],$$

where $O\big(\frac{1}{n}\big)$ is uniform with respect to $i=\overline{1,n}$.

Hence it follows that

$$Q_{ir} = \frac{1}{n^2 p(x_i)p(x_r)} K_N(x_i-x_r)\left[1+O\Big(\frac{1}{n}\Big)\right]. \tag{5}$$

Taking into account the relation

$$\max_{-\pi\le u\le\pi} |K_N(u)| = O(N) \tag{6}$$

and (5), we find

$$\sigma_{nN}^2 = \frac{\sigma^4}{2\pi^2(2N+1)^2n^2}\sum_{i=1}^n\sum_{j=1}^n K_N^2(x_i-x_j)\frac{1}{[p(x_i)p(x_j)]^2} + O\Big(\frac{1}{n}\Big). \tag{7}$$

Let $F(x)$ be a distribution function with density $p(x)$ and $F_n(x)$ be an empirical distribution function of the "sample" $x_1,x_2,\dots,x_n$, i.e., $F_n(x) = n^{-1}\sum_{k=1}^n I_{(-\infty,x)}(x_k)$, where $I_A(\cdot)$ is the indicator of the set A. Then the right side of (7) can be written as the integral

$$\sigma_{nN}^2 = \frac{\sigma^4}{2\pi^2(2N+1)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} K_N^2(t-s)\frac{dF_n(t)\,dF_n(s)}{[p(t)p(s)]^2} + O\Big(\frac{1}{n}\Big).$$

Further we have

$$\left|\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}K_N^2(t-s)\frac{dF_n(t)\,dF_n(s)}{[p(t)p(s)]^2}-\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}K_N^2(t-s)\frac{dF(t)\,dF(s)}{[p(t)p(s)]^2}\right|\le$$

$$\le I_1+I_2,$$

$$I_1=\left|\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}K_N^2(t-s)\frac{dF_n(s)}{[p(t)p(s)]^2}\big[dF_n(t)-dF(t)\big]\right|,$$

$$I_2=\left|\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}K_N^2(t-s)\frac{dF(t)}{[p(t)p(s)]^2}\big[dF_n(s)-dF(s)\big]\right|.$$

By integration by parts in the internal integral in I_1 we readily obtain

$$I_1\le 2\int_{-\pi}^{\pi}\frac{dF_n(s)}{p^2(s)}\int_{-\pi}^{\pi}\big|dF_n(t)-dF(t)\big|\,\big|\big(K_N'(t-s)p(t)-$$
$$-K_N(t-s)p'(t)\big)K_N(t-s)/p^3(t)\big|\,dt. \tag{8}$$

Since $\sup_{-\pi\le x\le\pi}|F_n(x)-F(x)|=O\big(\frac{1}{n}\big)$ and the relations [8][1]

$$\max_{-\pi\le u\le\pi}|K_N'(u)|=O(N^2),\quad \int_{-\pi}^{\pi}K_N^2(u)\,du=2N+1,$$
$$\int_{-\pi}^{\pi}|K_N(u)|\,du=O(\ln N) \tag{9}$$

are fulfilled, from (8) we have the estimate

$$I_1=O\left(\frac{N^2\ln N}{n}\right).$$

In the same manner we show that

$$I_2=O\left(\frac{N^2\ln N}{n}\right).$$

Therefore

$$(2N+1)\sigma_{nN}^2=\frac{\sigma^4}{4\pi^3}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\Phi_N(s-t)\frac{dt\,ds}{p(s)p(t)}+O\left(\frac{N\ln N}{n}\right), \tag{10}$$

where $\Phi_N(u)=\frac{2\pi}{2N+1}K_N^2(u)$ is the Fejér kernel.

We shall complete the definition of the function p^{-1} outside $[-\pi,\pi]$ as regards its periodicity and also note that $K_N(u)$ and $\Phi_N(u)$ are periodic functions with the period 2π. The continued function will be denoted by $g(x)$. Then

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\Phi_N(s-t)\frac{dt\,ds}{p(s)p(t)}=\int_{-\pi}^{\pi}p^{-2}(x)\,dx+\chi_n,$$

[1]See p. 115 in the Russian version of [8]: “Mir”, Moscow, 1965.

where

$$|\chi_n| \le \int_{-\pi}^{\pi} |\bar{\sigma}_N(x) - g(x)|dx,$$

$$\bar{\sigma}_N(x) = \int_{-\pi}^{\pi} \Phi_N(u) g(x-u) du.$$

Hence, on account of the theorem on convergence of the Fejér integral $\bar{\sigma}_N(x)$ to $g(x)$ in the norm of the space $L_1(-\pi,\pi)$ (see [9], p.481), we have $\chi_n \to 0$ for $n \to \infty$.

Therefore

$$(2N+1)\sigma_{nN}^2 \to \frac{\sigma^4}{4\pi^3} \int_{-\pi}^{\pi} p^{-2}(x)\, dx.$$

Now we shall prove (3). We have

$$D\mu_{nN}(x) = \sigma^2 \sum_{j=1}^{n} \frac{1}{np^2(x_j)} K_N^2(x-x_j) \left[1 + O\Big(\frac{1}{n}\Big)\right].$$

Applying the same reasoning as in deriving (10), we find

$$D\mu_{nN}(x) = \frac{\sigma^2}{n} \int_{-\pi}^{\pi} K_N^2(x-s) \frac{ds}{p(s)} + O\Big(\frac{N^2 \ln N}{n^2}\Big). \tag{11}$$

Therefore

$$EU_{nN} = \frac{\sigma^2}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \Phi_N(t-s) \frac{ds\, dt}{p(s)} + O\Big(\frac{N \ln N}{n}\Big) =$$
$$= \frac{\sigma^2}{(2\pi)^2} \int_{-\pi}^{\pi} p^{-1}(s)\, ds + O\Big(\frac{N \ln N}{n}\Big). \quad \square$$

Denote by the symbol $\xrightarrow{d}$ the convergence in distribution, and let ξ be a random variable having normal distribution with zero mean and variance 1.

Theorem 1. *Let x_i $i = \overline{1,n}$ be the same as in Lemma 2 and $\frac{N^2 \ln N}{n} \to 0$ for $n \to \infty$. Then, as n increases, $\sqrt{2N+1}(U_{nN} - \theta_1)\theta_2^{-1/2} \xrightarrow{d} \xi$.*

Proof. We have

$$\frac{U_{nN} - EU_{nN}}{\sigma_{nN}} = H_n^{(1)} + H_n^{(2)},$$

where

$$H_n^{(1)} = \sum_{j=1}^{n} \xi_j, \quad H_n^{(2)} = \frac{n}{2\pi(2N+1)\sigma_{nN}} \sum_{i=1}^{n} (\varepsilon_i^2 - E\varepsilon_i^2) Q_{ii}.$$

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

$$DH_n^{(2)} \le \frac{n^2 E\varepsilon_1^4}{(2\pi)^2(2N+1)^2\sigma_{nN}^2} \sum_{i=1}^n Q_{ii}^2 =$$
$$= \frac{E\varepsilon_1^4}{(2\pi)^2(2N+1)^2\sigma_{nN}^2 \cdot n^2} \sum_{i=1}^n \frac{1}{(p(x_i))^4} K_N^2(0) \left(1 + O\Big(\frac{1}{n}\Big)\right) \le$$
$$\le C \frac{1}{n\sigma_{nN}^2} = O\Big(\frac{N}{n}\Big),$$

whence $H_n^{(2)} \xrightarrow{P} 0$. Here and in what follows C is the positive constant varying from one formula to another and the letter P above the arrow denotes convergence in probability.

We will now prove that $H_n^{(1)} \xrightarrow{d} \xi$. To this end we will verify the validity of Corollaries 2 and 6 of Theorem 2 from [6]. We have to show whether the conditions contained in these statements are fulfilled for asymptotic normality of the square-integrable martingale-difference, which, by Lemma 1, is our sequence $\{\xi_k, \mathcal{F}_k\}_{k\ge1}$.

A direct calculation shows that $\sum_{k=1}^n E\xi_k^2 = 1$. Asymptotic normality will take place if for $n\to\infty$

$$\sum_{k=1}^n E\big[\xi_k^2 \cdot I(|\xi_k| \ge \varepsilon) \,|\, \mathcal{F}_{k-1}\big] \to 0 \tag{12}$$

and

$$\sum_{k=1}^n \xi_k^2 \xrightarrow{P} 1. \tag{13}$$

It is shown in [6] that the fulfillment of (13) and the condition $\sup\limits_{1\le k\le n} |\xi_k| \xrightarrow{P} 0$ implies the validity of (12) as well.

Since for $\varepsilon > 0$

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

to prove $H_n^{(1)} \xrightarrow{d} \xi$ we have to verify only (13) by the relation (15) to be given below.

We will establish $\sum_{k=1}^n \xi_k^2 \xrightarrow{P} 1$. For this it suffices to make sure that $E(\sum_{k=1}^n \xi_k^2 - 1))^2 \to 0$ for $n\to\infty$, i.e., due to $\sum_{i=1}^n E\xi_i^2 = 1$

$$E\Big(\sum_{k=1}^n \xi_k^2\Big)^2 = \sum_{k=1}^n E\xi_k^4 + 2 \sum_{1\le k_1<k_2\le n} E\xi_{k_1}^2 \xi_{k_2}^2 \to 1. \tag{14}$$

In the first place we find that $\sum_{k=1}^{n} E\xi_k^4 \to 0$ for $n \to \infty$. By virtue of the definitions of ξ_k and η_{ij} we write

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

where

$$L_n^{(1)} = \frac{n^4}{\pi^4(2N+1)^4\sigma_{nN}^4} E\varepsilon_1^4(E\varepsilon_1^4 - 3\sigma^4)\sum_{k=2}^{n}\sum_{j=1}^{k-1} Q_{jk}^4,$$

$$L_n^{(2)} = \frac{3n^4\sigma^4 E\varepsilon_1^4}{(2N+1)^4\sigma_{nN}^4\pi^4}\sum_{k=2}^{n}\Big(\sum_{j=1}^{k-1} Q_{jk}^2\Big)^2.$$

From (5) and (6) we obtain

$$|L_n^{(1)}| = C\frac{1}{n^4N^4\sigma_{nN}^4}\sum_{k=2}^{n}\sum_{j=1}^{k-1}\frac{K_N^4(x_j-x_k)}{[p(x_j)p(x_k)]^4}\Big[1+O\Big(\frac{1}{n}\Big)\Big] \le$$
$$\le Cn^{-2}\sigma_{nN}^{-4} = O\Big(\Big(\frac{N}{n}\Big)^2\Big)$$

and also

$$|L_n^{(2)}| = C\frac{1}{n^2N^4\sigma_{nN}^4}\sum_{k=2}^{n}\left(\frac{1}{n}\sum_{j=1}^{k-1}\frac{K_N^2(x_j-x_k)}{p(x_j)p(x_k)}\Big[1+O\Big(\frac{1}{n}\Big)\Big]\right)^2 \le$$
$$\le C\frac{1}{n^2N^4\sigma_{nN}^4}\sum_{k=1}^{n}\left(\frac{1}{n}\sum_{j=1}^{k-1}K_N^2(x_j-x_k)\right)^2 =$$
$$= C\frac{1}{n^2N^4\sigma_{nN}^4}\sum_{k=2}^{n}\Big(\int_{-\pi}^{\pi}\frac{K_N^2(x_k-u)}{p^2(u)}dF_n(u)\Big)^2 \le$$
$$\le C\frac{1}{n^2N^4\sigma_{nN}^4}\sum_{k=2}^{n}\Big\{\Big[K_N^2(x_k-u)p^{-1}(u)\,du\Big]^2 +$$
$$+\Big[\int_{-\pi}^{\pi}K_N^2(x_k-u)p^{-2}(u)\,d\big(F_n(u)-F(u)\big)\Big]^2\Big\}^2.$$

Hence, taking into account the relation (9) and the formula of integration by parts, we have $|L_n^{(2)}| = O\big(\frac{1}{n}\big)$. Therefore

$$\sum_{k=1}^{n} E\xi_k^4 \to 0 \quad \text{for } n \to \infty. \tag{15}$$

Let us now establish that $2\sum_{1\le k_1<k_2\le n} E\xi_{k_1}^2\xi_{k_2}^2 \to 1$ for $n\to\infty$. The definition of ξ_i implies

$$\xi_{k_1}^2\xi_{k_2}^2 = \Big(\sum_{i=1}^{k_1-1}\eta_{ik_1}^2\Big)\Big(\sum_{i=1}^{k_2-1}\eta_{ik_2}^2\Big) + \Big(\sum_{i=1}^{k_1-1}\eta_{ik_1}^2\Big)\Big(\sum_{i\ne s=1}^{k_2-1}\eta_{ik_2}\eta_{sk_2}\Big) +$$
$$+\Big(\sum_{i=1}^{k_2-1}\eta_{ik_2}^2\Big)\Big(\sum_{s\ne t=1}^{k_1-1}\eta_{sk_1}\eta_{tk_1}\Big) + \Big(\sum_{s\ne t=1}^{k_1-1}\eta_{sk_1}\eta_{tk_1}\Big)\Big(\sum_{k\ne r=1}^{k_2-1}\eta_{kk_1}\eta_{rk_2}\Big) =$$
$$= B_{k_1k_2}^{(1)} + B_{k_1k_2}^{(2)} + B_{k_1k_2}^{(3)} + B_{k_1k_2}^{(4)}.$$

Therefore

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

where

$$A_n^{(i)} = 2\sum_{1\le k_1<k_2\le n} EB_{k_1k_2}^{(i)}, \quad i=\overline{1,4}.$$

In the first place we consider $A_n^{(3)}$. By the definition of η_{ij} we obtain $E\eta_{ik_2}^2\eta_{sk_1}\eta_{tk_1}=0$, $s\ne t$, $k_1<k_2$. Thus

$$A_n^{(3)} = 0. \tag{16}$$

Let us derive an estimate of $A_n^{(2)}$. Divide the sum $EB_{k_1k_2}^{(2)}$ into two parts:

$$EB_{k_1k_2}^{(2)} = \sum_{i=1}^{k_1-1}\sum_{r\ne s=1}^{k_1} E\eta_{ik_1}^2\eta_{rk_2}\eta_{sk_2} + \sum_{i=1}^{k_1-1}\sum_{r\ne s=k_1+1}^{k_2-1} E\eta_{ik_1}^2\eta_{rk_2}\eta_{sk_2}.$$

The second term is equal to zero, since i cannot coincide with r or with s and $r\ne s$; in this case $E\eta_{ik_1}^2\eta_{rk_2}\eta_{sk_2}=0$, and $E\eta_{ik_1}^2\eta_{rk_2}\eta_{sk_2}=0$ also in the first term each time except for the case $s=k_1$ or $r=k_1$.

Thus

$$EB_{k_1k_2}^{(2)} = 2\sum_{i=1}^{k_1-1} E\big(\eta_{ik_1}^2\eta_{ik_2}\eta_{k_1k_2}\big).$$

Hence, using the definition of η_{ij} and the inequality $|Q_{ij}|\le C\frac{N}{n^2}$ obtained from (5) and (6), we find

$$\big|EB_{k_1k_2}^{(2)}\big| \le C\frac{1}{(2N+1)^2\sigma_{nN}^4}\sum_{i=1}^{k_1-1} Q_{ik_1}^2. \tag{17}$$

Next, taking into account statement (4) of Lemma 2 and the definition of σ^2_{nN}, from (17) we have

$$|A_n^{(2)}| \le C\frac{n}{N^2\sigma^4_{nN}}\sum_{k_1=2}^{n}\sum_{i=1}^{k_1-1}Q^2_{ik_1} \le C\frac{1}{n\sigma^2_{nN}} = O\Big(\frac{N}{n}\Big). \tag{18}$$

Consider now $A_n^{(4)}$. By the definition of η_{ij} we obtain

$$A_n^{(4)} = \frac{8n^4}{\pi^4(2N+1)^4\sigma^4_{nN}}\sum_{s<t<k_1<k_2}Q_{sk_1}Q_{sk_2}Q_{tk_1}Q_{tk_2} \le$$

$$\le C\frac{n^4}{N^4\sigma^4_{nN}}\Big[\Big|\sum_{s,t,k_1,k_2}Q_{sk_1}Q_{sk_2}Q_{tk_1}Q_{tk_2}\Big| +$$

$$+\Big|\sum_{k_1,s,t}Q^2_{k_1s}Q^2_{k_1t}\Big| + \Big|\sum_{k_1,s,t}Q_{k_1t}Q_{st}Q_{k_1s}Q_{ss}\Big|\Big] =$$

$$= C\frac{n^4}{N^4\sigma^4_{nN}}\big[|E_1| + |E_2| + |E_3|\big]. \tag{19}$$

According to (5) and (6) we write

$$E_1 = n^{-7}\sum_{s,t,k_1}K_N(x_s - x_{k_1})K_N(x_t - x_{k_1}) \times$$

$$\times \int_{-\pi}^{\pi}K_N(x_s - u)K_N(x_t - u)\,dF_n(u) + O\Big(\frac{N^2}{n}\Big).$$

Hence, integrating by parts and taking into account (9), we obtain

$$E_1 = n^{-7}\int_{-\pi}^{\pi}\sum_{s,t,k_1}K_N(x_s - x_{k_1})K_N(x_t - x_{k_1}) \times$$

$$K_N(x_s - u)K_N(x_t - u)p(u)\,du + O\Big(\frac{N^4\ln N}{n^5}\Big). \tag{20}$$

Applying the same operations three times, we represent (20) in the form

$$E_1 = n^{-4}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}K_N(z-u)K_N(z-t)K_N(y-u)K_N(y-t) \times$$

$$\times p(y)p(u)p(z)p(t)\,du\,dt\,dz\,dy + O\Big(\frac{N^4\ln N}{n^5}\Big) =$$

$$= O\Big(\frac{N\ln^3 N}{n^4}\Big) + O\Big(\frac{N^4\ln N}{n^5}\Big).$$

Thus

$$\frac{n^4}{N^4\sigma_{nN}^4}|E_1| = O\Big(\frac{\ln^3 N}{N}\Big) + O\Big(\frac{N^2\ln N}{n}\Big). \tag{21}$$

Further, it is not difficult to show

$$\begin{aligned}\frac{n^4}{N^4\sigma_{nN}^4}|E_2| &= O\Big(\frac{N^2}{n}\Big),\\ \frac{n^4}{N^4\sigma_{nN}^4}|E_3| &= O\Big(\frac{N^2}{n}\Big).\end{aligned} \tag{22}$$

Therefore (19), (21), and (22) imply

$$A_n^{(4)} = O\Big(\frac{N^2\ln N}{n}\Big) + O\Big(\frac{\ln^3 N}{N}\Big). \tag{23}$$

Finally, we will show that $A_n^{(1)} \to 1$ for $n\to\infty$. For this represent $A_n^{(1)}$ in the form $A_n^{(1)} = Q_n^{(1)} + Q_n^{(2)}$, where

$$Q_n^{(1)} = 2\sum_{k_1<k_2}\Big(\sum_{i=1}^{k_1-1} E\eta_{ik_1}^2\Big)\Big(\sum_{j=1}^{k_2-1} E\eta_{ik_2}^2\Big),$$

$$Q_n^{(2)} = 2\Big(\sum_{k_1<k_2} EB_{k_1k_2}^{(1)} - \sum_{k_1<k_2}\Big(\sum_{i=1}^{k_1-1} E\eta_{ik_1}^2\Big)\Big(\sum_{j=1}^{k_2-1} E\eta_{ik_2}^2\Big)\Big).$$

From the definition of σ_{nN}^2 it follows that

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

where

$$\begin{aligned}\sum_{k=2}^{n}\Big(\sum_{i=1}^{k-1} E\eta_{ik}^2\Big)^2 &\le C\frac{n^4}{N^4\sigma_{nN}^4}\sum_{k=2}^{n}\Big(\sum_{i=1}^{k-1} Q_{ik}^2\Big)^2 \le\\ &\le C\frac{1}{n\sigma_{nN}^4} = O\Big(\frac{N^2}{n}\Big).\end{aligned}$$

Therefore

$$Q_n^{(1)} = 1 + O\big(N^2/n\big). \tag{24}$$

Let us now show that $Q_n^{(2)} \to 0$. $Q_n^{(2)}$ can be written as

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

But

$$E\eta_{ik_1}^2\eta_{ik_2}^2 \le C\frac{n^4}{N^4\sigma_{nN}^4}Q_{ik_1}^2\cdot Q_{ik_2}^2 \le$$
$$\le C\frac{1}{n^4N^4\sigma_{nN}^4}\Big(\max_{-\pi\le u\le\pi}|K_N(u)|\Big)^4 = O\Big(\frac{1}{n^4\sigma_{nN}^4}\Big).$$

Similarly, $E\eta_{ij}^2 = O\big(n^{-2}\sigma_{nN}^{-2}\big)$. Therefore

$$\operatorname{cov}(\eta_{ik_1}^2,\eta_{ik_2}^2) = O\Big(\frac{1}{n^4\sigma_{nN}^4}\Big). \tag{25}$$

Further, since $\sum_{1\le k_1<k_2\le n}(k_1-1) = O(n^3)$, (25) implies

$$Q_n^{(2)} = O\Big(\frac{N^2}{n}\Big). \tag{26}$$

Thus, according to (24) and (26)

$$A_n^{(1)} = 1 + O\big(N^2/n\big). \tag{27}$$

Combining the relations (16), (18), (23) and (27), we finally obtain

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

Therefore

$$\frac{U_{nN}-EU_{nN}}{\sigma_{nN}} \xrightarrow{d} \xi.$$

Further, due to Lemma 2, $EU_{nN} = \theta_1 + O(\frac{N\ln N}{n})$ and $(2N+1)\sigma_{nN}^2 \to \theta_2$, and hence we obtain

$$(2N+1)^{1/2}(U_{nN}-\theta_1)\theta_2^{-1/2} \xrightarrow{d} \xi. \quad \square$$

Denote

$$T_{nN} = \frac{n}{2\pi(2N+1)}\int_{-\pi}^{\pi}\big[\mu_{nN}(x)-\mu(x)\big]^2\,dx.$$

Theorem 2. *Let x_i, $i=\overline{1,n}$, be the same as in Lemma 2 and the function $\mu(x)$ with period 2π have bounded derivatives up to the second order. Moreover, if $N^2\ln N/n \to 0$ and $n\ln^2 N/N^{9/2} \to 0$ for $n\to\infty$, then $\sqrt{2N+1}(T_{nN}-\theta_1)\theta_2^{-1/2} \xrightarrow{d} \xi$.*

Before we proceed to proving the theorem, we have to show

$$\int_{-\pi}^{\pi} |K_N'(u)|du = O(N \ln N). \tag{28}$$

Denote $\widetilde{D}_\nu(u) = \sum_{k=1}^{\nu} \sin ku$. Then by virtue of the Abel transformation we have

$$K_N'(u) = -\sum_{k=1}^{N} k \sin ku = \sum_{\nu=1}^{N-1} \widetilde{D}_\nu(u) + N\widetilde{D}_N.$$

It is well known [8] that $\eta_\nu = (\ln \nu)^{-1} \int_{-\pi}^{\pi} |\widetilde{D}_\nu(u)|du \to 1$ for $\nu \to \infty$. Denote $b_N = \sum_{\nu=1}^{N-1} \ln \nu$. Then by the Toeplitz lemma

$$R_N = \frac{1}{b_N} \sum_{\nu=1}^{N-1} \ln \nu \cdot \eta_\nu \to 1.$$

Therefore

$$\int_{-\pi}^{\pi} |K_N'(u)|du \le \sum_{\nu=1}^{N-1} \int_{-\pi}^{\pi} |\widetilde{D}_\nu(u)|du + N \int_{-\pi}^{\pi} |\widetilde{D}_N(u)|du =$$

$$= b_N \cdot R_N + N \int_{-\pi}^{\pi} |\widetilde{D}_N(u)|du = O(N \ln N).$$

Let us return to the proof of the theorem. We have

$$T_{nN} = U_{nN} + A_{1n} + A_{2n},$$

$$A_{1n} = \frac{n}{\pi(2N+1)} \int_{-\pi}^{\pi} \big[\mu_{nN}(x) - E\mu_{nN}(x)\big]\big[E\mu_{nN}(x) - \mu(x)\big]dx,$$

$$A_{2n} = \frac{n}{2\pi(2N+1)} \int_{-\pi}^{\pi} \big[E\mu_{nN}(x) - \mu(x)\big]^2 dx.$$

It is not difficult to find

$$\sqrt{2N+1}E|A_{1n}| \le \frac{n\sigma^2}{2\pi\sqrt{2N+1}} \Big(\sum_{j=1}^{n} \Delta_j^2 \Big[\int_{-\pi}^{\pi} K_n(y - x_j) \times$$

$$\times \big(E\mu_{nN}(y) - \mu(x)\big) dy\Big]^2 \Big)^{1/2}.$$

But

$$E\mu_{nN}(y) = \int_{-\pi}^{\pi} \mu(x) \frac{1}{p(x)} K_N(y - x)\, dF_n(x) \Big(1 + O\big(\frac{1}{n}\big)\Big)$$

and

$$\int_{-\pi}^{\pi} \mu(x)p^{-1}(x)K_N(y-x)\,dF_n(x) =$$
$$= \int_{-\pi}^{\pi} \mu(x)K_N(y-x) + O\Big(\frac{1}{n}\int_{-\pi}^{\pi}|K'_N(u)|du\Big).$$

It is well known ([10], p.22) that

$$\int_{-\pi}^{\pi} \mu(x)K_N(y-x)\,dx = \mu(y) + O\Big(\frac{\ln N}{N^2}\Big)$$

uniformly in $y \in [-\pi,\pi]$. By virtue of (28) this gives us

$$E\mu_{nN}(x) = \mu(x) + O\Big(\frac{\ln N}{N^2}\Big) + O\Big(\frac{N\ln N}{n}\Big). \tag{29}$$

Therefore

$$\sqrt{2N+1}E|A_{1n}| \le C\Big[\Big(\frac{n\ln^2 N}{N^{9/2}}\Big)^{1/2}\frac{\ln N}{N^{1/4}} + $$
$$+\Big(\frac{N^2\ln N}{n}\Big)^{1/2}\frac{\ln^{3/2}N}{\sqrt{N}}\Big] \to 0. \tag{30}$$

Further, from (29) we have

$$\sqrt{2N+1}A_{2n} \le C\Big(\frac{n\ln^2 N}{N^{9/2}} + \frac{N^2}{n}\,\frac{\ln^2 N}{\sqrt{N}}\Big) \to 0. \tag{31}$$

Finally, the statement of Theorem 2 directly follows from Theorem 1, (30), and (31).

Using Theorems 1 and 2, it is easy to solve the problem concerning testing of the hypothesis on $\mu(x)$. Given σ^2, it is required to verify the hypothesis $H_0 : \mu(x) = \mu_0(x)$. The critical region is defined approximately by the inequality $U_{nN} \ge d_n(\alpha)$ or $T_{nN} \ge d_n(\alpha)$, where

$$d_n(\alpha) = \sigma^2\big(L_1 + (2N+1)^{-1/2}L_2\big)\lambda_\alpha,$$
$$L_1 = ((2\pi)^{-2}\int_{-\pi}^{\pi} p^{-1}(x)\,dx, \quad L_2 = \Big(\frac{1}{4\pi^3}\int_{-\pi}^{\pi} p^{-2}(x)\,dx\Big)^{1/2},$$

and λ_α is the quantile of level α of standard normal distribution.

Let now σ^2 be unknown. We call an $\sqrt{N}$-consistent estimate of variance σ^2, for instance,

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}\big(Y_i - \mu_{n\lambda}(x_i)\big)^2,$$

where $\lambda = \lambda(n) \to \infty$ is a sequence such that $\frac{\lambda}{N} \to 0$, $\frac{N\ln^2\lambda}{\lambda^4} \to 0$ and $\frac{N\lambda^4}{n} \to 0$ for $n \to \infty$.

Indeed, using the expressions (11) and (29), we easily find

$$\sqrt{N}(ES_n^2-\sigma^2)=O\Big(\Big(\frac{N\lambda}{n}\Big)^{1/2}\Big)+O\Big(\frac{N^{1/2}\ln\lambda}{\lambda^2}\Big). \tag{32}$$

Denote

$$Z_j=Y_j-R_j,$$
$$R_j=\sum_{k=1}^{n}Y_k\Delta_k K_\lambda(x_j-x_k).$$

Then

$$n^2DS_n^2=\sum_{j=1}^{n}DZ_j^2+\sum_{i\neq i_1}\operatorname{cov}(Z_i^2,Z_{i_1}^2).$$

Simple calculations show that $\operatorname{cov}(Z_j^2,Z_{j_1}^2)=O\Big(\frac{\lambda^4}{n}\Big)$. Therefore $DS_n^2=O\Big(\frac{\lambda^4}{n}\Big)$. This and (32) imply $\sqrt{N}(S_n^2-\sigma^2)\xrightarrow{P}0$.

Corollary. *Let the conditions of Theorem 2 be fulfilled. Moreover, let* $\frac{\lambda}{n}\to 0$, $\frac{N\lambda^4}{n}\to 0$ *and* $\frac{N\ln^2\lambda}{\lambda^4}\to 0$. *Then*

$$S_n^{-2}L_2^{-1}\sqrt{2N+1}(U_{nN}-S_n^2L_1)\xrightarrow{d}\xi,$$
$$S_n^{-2}L_2^{-1}\sqrt{2N+1}(T_{nN}-S_n^2L_1)\xrightarrow{d}\xi.$$

This corollary enables one to construct a test for verifying $H_0:\mu(x)=\mu_0(x)$. The critical region is defined approximately by the inequality $U_{nN}\geq\widetilde{d}_n(\alpha)$ or $T_{nN}\geq\widetilde{d}_n(\alpha)$, where $\widetilde{d}_n(\alpha)$ is obtained from $d_n(\alpha)$ by using S_n^2 instead of σ^2.

Consider now the local behavior of the test power in the case where the critical region is of the form $\{x\in R^1,\ x\geq d_n(\alpha)\}$. More exactly, find a distribution of the quadratic functional U_{nN} under a sequence of alternatives close to the hypothesis $H_0:\mu(x)=\mu_0(x)$. The sequence is written as

$$H_1:\bar{\mu}(x)=\mu_0(x)+\gamma_n\varphi(x)+o(\gamma_n), \tag{33}$$

where $\gamma_n\to 0$ appropriately and $o(\gamma_n)$ is uniform in $x\in[-\pi,\pi]$.

Theorem 3. *Let* $\bar{\mu}_n(x)$ *satisfy the conditions of Theorem 2. If* $2N+1=n^\delta$, $\gamma_n=n^{-1/2+\delta/4}$, $\frac{2}{9}<\delta<\frac{1}{2}$, *then under the alternative* H_1 *the statistic* $(2N+1)^{1/2}(U_{nN}-\theta_1)$ *is distributed in the limit normally* $\big(\frac{1}{2\pi}\int_{-\pi}^{\pi}\varphi^2(u)du,\sqrt{\theta_2}\big)$.

Proof. Let us represent U_{nN} as the sum

$$U_{nN} = \frac{n}{2\pi(2N+1)} \int_{-\pi}^{\pi} \big(\mu_{nN}(x) - E_1\mu_{nN}(x)\big)^2 dx + \\ + \frac{n}{\pi(2N+1)}\gamma_n \int_{-\pi}^{\pi} \big[\mu_{nN}(x) - E_1\mu_{nN}(x)\big]\widetilde{\varphi}_n(x)\,dx + \\ + \frac{n}{2\pi(2N+1)}\gamma_n^2 \int_{-\pi}^{\pi} \widetilde{\varphi}_n^2(x)\,dx = A_1(n) + A_2(n) + A_3(n),$$

where $E_1(\cdot)$ denotes the mathematical expectation under the hypothesis H_1,

$$\widetilde{\varphi}_n(x) = \sum_{j=1}^{n} \varphi(x_j)\Delta_j K_n(x - x_j).$$

Due to Theorem 1 one can readily assertain that $\sqrt{2N+1}(A_1(n) - \theta_1)$ is distributed asymptotically normal $(0, \sqrt{\theta}_2)$.

By analogy with the proof of Lemma 2 we find

$$\sqrt{2N+1}A_3(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Big(\int_{-\pi}^{\pi} \varphi(y) K_N(x-y) dy \Big)^2 dx + O\Big(\frac{N^2 \ln N}{n}\Big).$$

Hence, by virtue of theorem 2 from [9], p.474, we have

$$\sqrt{2N+1}A_3(n) \to \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi^2(u) du.$$

Further, for our choice of N and γ_n we can show by simple calculations that

$$\sqrt{2N+1}E|A_2(n)| \le C\Big(\frac{\ln^2 n}{n^{\delta/4}} + \frac{\ln n}{n^{1-7\delta/4}}\Big). \quad \square$$

Thus the local behaviour of the power $P_{H_1}(U_{nN} \ge d_n(\alpha))$ is

$$P_{H_1}\big(U_{nN} \ge d_n(\alpha)\big) \to 1 - \Phi\Big(\lambda_\alpha - \theta_2^{-1/2} \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi^2(u)\,du\Big). \qquad (34)$$

Since $\int_{-\pi}^{\pi} \varphi^2(u) du > 0$ and is equal to zero iff $\varphi(x) = 0$, from (34) we conclude that the test for the hypothesis $H_0 : \mu(x) = \mu_0(x)$ against alternatives of the form (33) is asymptotically strictly unbiased.

Remark. Similar results can be obtained by the same method for the kernel estimator of Priestley and Chao [1].

References

1. M.B. Priestley and M.T. Chao, Nonparametric function fitting. *J. Roy. Statist. Asoc. ser. B* **34**(1972), 385-392.

2. N.N. Chentsov, Estimation of the unknown distribution density by observations. (Russian) *Dokl. Akad. Nauk SSSR* **147**(1962), 45-48.

3. N.B. Smirnov and I.V. Dunin-Barkovskii, A course in probability theory and mathematical statistics for technical applications. (Russian) *Nauka, Moscow,* 1969.

4. B.T. Polyak and A.V. Tsibakov, Asymptotic optimality of the C_p-test in the projection estimation of the regression. (Russian) *Teor. veroyatnost. i primenen.* **35**(1990), No. 2, 305-317.

5. R.L. Eubank, J.D. Hart, and P. Speckman, Trigonometric series regression estimators with application to partially linear models. *J. Multivariate Anal.* **32**(1990), 70-83.

6. R.Sh. Liptser and A.N. Shiryayev, A functional central limit theorem for semimartingales. (Russian) *Teor. veroyatnost. i primenen.* **25**(1980), No.4, 683-703.

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

8. A. Zygmund, Trigonometric series, vol. 1. *Cambridge University Press, Cambridge,* 1959.

9. A.N. Kolmogorov and S.V. Fomin, Elements of the theory of functions and functional analysis. (Russian) *"Nauka", Moscow,* 1989.

10. D. Jackson, The theory of approximation. *American Mathematical Society, New York,* 1930.

(Received 1.07.1992)

Author's address:
Faculty of Mechanics and Mathamatics
I.Javakhishvili Tbilisi State University
2 University St., 380043 Tbilisi
Republic of Georgia