Ann. Inst. Statist. Math.
31 (1979), Part A, 279-288

STRONG CONSISTENCY OF DENSITY ESTIMATION BY ORTHOGONAL SERIES METHODS FOR DEPENDENT VARIABLES WITH APPLICATIONS

IBRAHIM A. AHMAD*

(Received Jan. 17, 1978; revised Apr. 21, 1979)

Abstract

Among several widely used methods of nonparametric density estimation is the technique of orthogonal series advocated by several authors. For such estimate when the observations are assumed to have been taken from strong mixing sequence in the sense of Rosenblatt [7] we study strong consistency by developing probability inequality for bounded strongly mixing random variables. The results obtained are then applied to two estimates of the functional $\Delta(f)=\int f^2(x)dx$ where strong consistency is established. One of the suggested two estimates of $\Delta(f)$ was recently studied by Schuler and Wolff [8] in the case of independent and identically distributed observations where they established consistency in the second mean of the estimate.

1. Introduction

Among the widely used methods of nonparametric density estimation is the method of orthogonal series. Several authors discussed this method, among them Cencov [2], Kronmal and Tarter [6], Schwartz [9], and Watson [10]. Assume that $f(x)$ is a square integrable probability density function (p.d.f.). Thus $f(x)$ can be expanded by orthogonal series, viz.,

$$f(x)=\sum_{j=0}^{\infty}\theta_j\phi_j(x) , \tag{1.1}$$

* Research supported in part by the National Research Council of Canada and in part by McMaster University Research Board. Now at Memphis State University, Memphis, Tennessee 38152, U.S.A.

AMS subject classification: Primary: 62G05; Secondary: 60G10.
Key words and phrases: Density estimation, orthogonal series, strong mixing sequences, strong consistency, non-parametric inference.

where $\theta_j = \int f(x)\phi_j(x)dx$, $j=1, 2, \cdots$. We need the following definitions:

DEFINITION 1.1. Let $\{X_n\}$ be a sequence of random variables and let $\mathcal{F}(m, n)$ denote the σ-field generated by $X_{m+1}, \cdots, X_{m+n}$ for all $m, n \geqq 1$. Let $A \in \mathcal{F}(1, m)$ and $B \in \mathcal{F}(m+n, \infty)$ and let α be non-negative function of positive integers such that $\alpha(n) \to 0$ as $n \to \infty$. Then $\{X_n\}$ is said to be strong mixing sequence with mixing numbers $\alpha(n)$ if

$$|\mathrm{P}(AB) - \mathrm{P}(A)\,\mathrm{P}(B)| \leqq \alpha(n) .$$

DEFINITION 1.2. A sequence $\{X_n\}$ of random variables is said to be stationary (in the strict sence) if the joint distribution of $(X_{n_1+k}, X_{n_2+k}, \cdots, X_{n_s+k})$ does not depend on k, for any integer k.

Remark 1.1. If $\{X_n\}$ is a stationary strong mixing sequence of random variables with mixing numbers $\alpha(n)$ and if we set $\eta_{N_n} = \phi_N(X_n)$ for a sequence of functions $\{\phi_N\}$ defined on the real line, then the array $\{\eta_{N_n}\}$ is also stationary row-wise strong mixing with mixing numbers $\alpha(n)$; see Section 18.5 of Ibragimov and Linnik [5].

Strong mixing sequences of random variables include many other dependent forms as special cases and the reader is referred to Ibragimov and Linnik [5] for further details.

Suppose that $\{X_n\}$ is a strong mixing stationary sequence of random variables with marginal p.d.f. $f(x)$ which is square integrable. Let $X_1, \cdots, X_n$ denote the first n observation in the sequence and let $F_n(x) = n^{-1} \sum_{j=1}^{n} I(x, X_j)$ denote the marginal empirical distribution function (d.f.). An unbiased estimate of θ_j, $j \geqq 1$ is given by

$$\hat{\theta}_j = n^{-1} \sum_{l=1}^{n} \phi_j(X_l) . \tag{1.2}$$

Consider the estimator of $f(x)$ given by,

$$\hat{f}(x) = \sum_{i=0}^{q(n)} \hat{\theta}_i \phi_i(x) , \tag{1.3}$$

where $q(n)$ is an integer-valued function of n such that $q(n) \to \infty$ as $n \to \infty$.

In Section 2 a basic lemma on probability inequalities for stationary strong mixing bounded random variables is presented. The result is an extension to dependent variables of Theorem 2 of Hoeffding [4]. In Section 3 we give conditions under which $\sup_x |\hat{f}(x) - f(x)|$ and $\int [\hat{f}(x) - f(x)]^2 dx$ converge to 0 with probability one (w.p. 1) as $n \to \infty$, and reduction of these results to the independent case is discussed. Finally in Section 4, applications of the results of Section 2 to estimating $\Delta(f) = \int f^2(x)dx$ are presented.

2. Basic lemma

The following lemma is an extension of Theorem 2 of Hoeffding [4] to stationary strong mixing r.v.'s.

LEMMA 2.1. *Let $\{X_n\}$ be a stationary strong mixing random variables such that $|X_n| \leqq C$, w.p. 1, and $\mathrm{E}\, X_n = \mu_n$, $n=1, 2, \cdots$. Set $\mu = (1/n) \sum_{j=1}^{n} \mu_j$, and let $m = m(n)$ and $p = p(n)$ be two positive integers such that $m \to \infty$ and $p \to \infty$ as $n \to \infty$. Then for all $t \geqq 0$, and all $n \geqq 1$,*

$$\mathrm{P}[\bar{X} - \mu \geqq t] \leqq 2\{\exp(-pt^2/2C^2)\}[1 + K\alpha(m)]^p, \tag{2.1}$$

where $K = 2\sqrt{e}$.

PROOF. Let m and p be such that $n = 2mp$. If this is not the case the proof may be easily altered but the result remains valid. Set

$$U_j = \sum_{i=(2j-2)m+1}^{(2j-1)m} X_i$$

and

$$V_j = \sum_{i=(2j-1)m+1}^{2jm} X_i ,$$

$j = 1, 2, \cdots, p$. Note that with probability one $|U_j| \leqq mC$ and $|V_j| \leqq mC$, $j = 1, 2, \cdots, p$. Set $\theta_j = \mathrm{E}\, U_j$ and $\gamma_j = \mathrm{E}\, V_j$, $j = 1, 2, \cdots, p$. Then we have

$$\begin{aligned} \mathrm{P}[\bar{X} - \mu \geqq t] &= \mathrm{P}\left[\sum_{i=1}^{n} X_i - n\mu \geqq nt\right] \\ &= \mathrm{P}\left[\sum_{j=1}^{p} (U_j - \theta_j) + \sum_{j=1}^{p} (V_j - \gamma_j) \geqq nt\right] \\ &\leqq \mathrm{P}\left[\sum_{j=1}^{p} (U_j - \theta_j) \geqq mpt\right] + \mathrm{P}\left[\sum_{j=1}^{p} (V_j - \gamma_j) \geqq mpt\right] . \end{aligned} \tag{2.2}$$

The lemma will be proved if we show that each of the two terms of the last upper bound of (2.2) is bounded above by $\{\exp(-pt^2/2C)\}[1 + K\alpha(m)]^p$. Set $\theta = p^{-1} \sum_{j=1}^{p} \theta_j$. Thus,

$$\mathrm{P}\left[\sum_{j=1}^{p} (U_j - \theta_j) \geqq mpt\right] \leqq [\exp\{-h(mpt + p\theta)\}]\, \mathrm{E}\left\{\exp\left(h \sum_{j=1}^{p} U_j\right)\right\} .$$

But as in Hoeffding [4], Theorem 2

$$\mathrm{E}\prod_{j=1}^{p} e^{h(U_j-\theta_j)} \leqq e^{-h\theta_j}\left[\frac{mC-\theta_j}{2Cm}e^{-hCm}+\frac{\theta_j+Cm}{2Cm}e^{jCm}\right] \tag{2.3}$$
$$=e^{-h_j p_j}(1-p_j+p_j e^{h_j}) ,$$

where $p_j=(\theta_j+Cm)/2Cm$ and $h_j=2hCm$, $j=1,\cdots,p$. Let

$$f(h_j)=-h_j p_j+\log(1-p_j+p_j e^{h_j}) . \tag{2.4}$$

Then as in Hoeffding [4], Theorem 2, we get

$$\mathrm{E}\, e^{h(U_j-\theta_j)} \leqq e^{(hCm)^2/2} . \tag{2.5}$$

Now,

$$I=\mathrm{E}\prod_{j=1}^{p} e^{h(U_j-\theta_j)}=\mathrm{E}\left\{\prod_{j=1}^{p-1} e^{h(U_j-\theta_j)}[\mathrm{E}\,(e^{h(U_p-\theta_p)}\,|\,U_1,\cdots,U_{p-1}) \right. \tag{2.6}$$
$$\left. -\mathrm{E}\, e^{h(U_p-\theta_p)}]\right\}+\mathrm{E}\prod_{j=1}^{p-1} e^{h(U_j-\theta_j)}\,\mathrm{E}\, e^{h(U_p-\theta_p)} .$$

But $|e^{hU_j}|\leqq e^{h|U_j|}\leqq e^{hmC}$ with probability one, thus $e^{hU_j-h\theta_j}\leqq e^{hmC}$ with probability one. By Lemma 5.2 of Dvoretzky [3] we have that the RHS of (2.6) is no greater than

$$\mathrm{E}\left\{\prod_{j=1}^{p-1} e^{h(U_j-\theta_j)}[2e^{hmC}\alpha(m)]\right\}+\mathrm{E}\prod_{j=1}^{p-1} e^{h(U_j-\theta_j)}e^{(hCm)^2/2} \tag{2.7}$$
$$=\mathrm{E}\prod_{j=1}^{p-1} e^{h(U_j-\theta_j)}\{2e^{hmC}\alpha(m)+e^{(hCm)^2/2}\}$$
$$\leqq e^{(hCm)^2p/2}\{1+\alpha(m)[2e^{hmC-(hCm)^2/2}]\}^p$$

where the last inequality is obtained by successive application of Lemma 5.2 of Dvoretzky [3] to $U_{p-1},\cdots,U_1$. Since

$$2e^{hmc-(hmC)^2/2}=2\sqrt{e}\cdot e^{-1/2(1-hmC)^2}\leqq 2\sqrt{e}=K ,$$

say, then we arrive at,

$$\mathrm{E}\prod_{j=1}^{p} e^{h(U_j-\theta_j)}\leqq e^{(hmC)^2p/2}[1+\alpha(m)]^p . \tag{2.8}$$

Hence

$$\mathrm{P}\left[\sum_{j=1}^{p}(U_j-\theta_j)\geqq mpt\right]\leqq e^{-hmpt+(hmc)^2p/2}[1+\alpha(m)K]^p , \tag{2.9}$$

where $K=2\sqrt{e}$. Let

$$f(h)=-hmpt+\frac{(hmC)^2}{2}p ,$$

thus $f(h)$ is minimized at $h=t/mC^2$ and thus

$$\text{(2.10)} \qquad \mathrm{P}\left[\sum_{j=1}^{p}(U_j-\theta_j)\geqq mpt\right]\leqq\{\exp(-pt^2/2C^2)\}\{1+K\alpha(m)\}^p .$$

Similarly we can establish that for all $t\geqq 0$,

$$\text{(2.11)} \qquad \mathrm{P}\left[\sum_{j=1}^{p}(V_j-\gamma_j)\geqq mpt\right]\leqq\{\exp(-pt^2/2C^2)\}\{1+K\alpha(m)\}^p .$$

Hence (2.1) follows from (2.10) and (2.11).

Remark. When the X_n's are independent $\alpha(n)=0$ for all n, thus in the above Lemma 2.1 if we take $p(n)=n$ for all n then it reduces to Theorem 2 of Hoeffding [4].

The next Corollary gives an indication of some possible choices of m and p.

COROLLARY 2.1. *Let $\lambda>0$ be real number and let $[x]$ denote the largest integer less than or equal to x. Then for all $t\geqq 0$,*

$$\text{(2.12)} \qquad \mathrm{P}[\bar{X}-\mu\geqq t]\leqq 2[\exp\{-[n^{(1-\lambda)}]t^2/2C^2\}]\{1+K\alpha([n^\lambda])\}^{[n^{(1-\lambda)}]} .$$

PROOF. In the above lemma take $m=[n^\lambda]$ and $p=[n^{(1-\lambda)}]$.

3. Strong consistency of $\hat{f}(x)$

In this section we shall use Lemma 2.1 and its corollary to give sufficient conditions for strong consistency of $\hat{f}(x)$. The main results are Theorems 3.1 and 3.2. Assume that $|\phi_j(x)|\leqq C$ for all $j\geqq 0$.

THEOREM 3.1. *If $m(n)$, $p(n)$ and $q(n)$ are integer-valued functions such that for any $\varepsilon>0$,*

$$\sum_{n=1}^{\infty} e^{-\varepsilon p(n)/q^2(n)}[1+K\alpha(m(n))]^{p(n)}<\infty \text{ , } \textit{then}$$

$$\text{(3.1)} \qquad \sup_x|\hat{f}(x)-f(x)|\to 0 \quad \textit{w.p. } 1 \textit{ as } n\to\infty .$$

PROOF. Note that

$$\sup_x|\hat{f}(x)-f(x)|\leqq\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|+\sup_x|\mathrm{E}\,\hat{f}(x)-f(x)| .$$

But

$$\text{(3.2)} \qquad \sup_x|\mathrm{E}\,\hat{f}(x)-f(x)|=\sup_x\left|\sum_{i=1+q(n)}^{\infty}\theta_i\phi_i(x)\right|\to 0 \qquad \text{as } n\to\infty .$$

Thus we need only to show that $\sup|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\to 0$ w.p. 1 as $n\to\infty$. To this end let

$$\Phi_n(x, y)=\sum_{j=0}^{q(n)} \phi_j(x)\phi_j(y) . \tag{3.3}$$

Then $\hat{f}(x)$ can be written as follows:

$$\hat{f}(x)=n^{-1}\sum_{i=1}^{n} \Phi_n(x, X_i) . \tag{3.4}$$

Now it follows from Remark 1.1 that random variables $\{\Phi_n(x, X_i)\}_{i=1}^n$ are stationary strong mixing such that

$$|\Phi_n(x, X_i)|=\left|\sum_{j=0}^{q(n)} \phi_j(x)\phi_j(X_i)\right| \leqq q(n)C^2 .$$

Thus, for any $\varepsilon_n \geqq 0$,

$$\begin{aligned}
(3.5)\quad & \mathrm{P}\,[\sup_x |\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\geqq\varepsilon_n] \\
&= \mathrm{P}\left[\sup_x |[n(q(n)+1)]^{-1}\sum_{j=1}^{n}\Phi_n(x, X_i)-(q(n)+1)^{-1}\,\mathrm{E}\,\Phi_n(x, X_1)|\right. \\
&\qquad \left.\geqq \varepsilon_n(q(n)+1)^{-1}\right] \\
&\leqq \mathrm{P}\left[(q(n)+1)^{-1}\sum_{j=1}^{n}\sup_x|\Phi_n(x, X_i)-\mathrm{E}\,\Phi_n(x, X_1)|\geqq\varepsilon_n(q(n)+1)^{-1}\right] \\
&\leqq \mathrm{P}\left[(q(n)+1)^{-1}\sum_{i=1}^{n}\sum_{j=0}^{q(n)}|\phi_j(X_i)-\mathrm{E}\,\phi_j(X_1)|\geqq\varepsilon_n n/C(q(n)+1)\right] .
\end{aligned}$$

Let

$$Y_{in}=[q(n)+1]^{-1}\sum_{j=0}^{q(n)}|\phi_j(X_i)-\mathrm{E}\,\phi_j(X_1)|$$

and set $\mu_{in}=\mathrm{E}\,Y_{in}$, $i=1, 2,\cdots, n$. Note that $\mu_{n1}=\cdots=\mu_{nn}$ and hence

$$\mathrm{P}\,[\sup_x |\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\geqq\varepsilon_n]\leqq\mathrm{P}\left[n^{-1}\sum_{i=1}^{n}(Y_{in}-\mu_{1n})\geqq\frac{\varepsilon_n n}{C(q(n)+1)}-\mu_{1n}\right] . \tag{3.6}$$

Choosing $\varepsilon_n=\varepsilon+C(q(n)+1)\mu_{1n}/n$ for any $\varepsilon>0$ we get that the upper bound of (3.6) is equal to

$$\mathrm{P}\left[n^{-1}\sum_{j=1}^{n}(Y_{in}-\mu_{1n})\geqq n\varepsilon/C(q(n)+1)\right] .$$

Now, it follows from Remark 1.1 that $\{Y_{in}\}_{i=1}^n$ are stationary mixing and hence by Lemma 2.1, the above upper bound is majorized by

$$2\{\exp(-p(n)t^2(n)/4C^2)\}\{1+K\alpha(n)\}^{p(n)} , \tag{3.7}$$

where $t(n)=\varepsilon n/C(q(n)+1)$. The Borel-Centelli lemma yields the result.

COROLLARY 3.1. *If for some real number $\lambda>0$, and for all $\varepsilon>0$,*

$$\sum_{n=1}^{\infty} e^{-\varepsilon[n^{(1-\lambda)}]/(q(n))^2}(1+K\alpha([n^\lambda]))^{[n^{(1-\lambda)}]}<\infty ,$$

then (3.1) *holds.*

PROOF. In Theorem 3.1 take $m(n)=[n^\lambda]$ and $p(n)=[n^{(1-\lambda)}]$.

Remark. If $\{X_i\}$ are independent identically distributed then in view of the remark following Lemma 2.1 we only need to assume that for any $\varepsilon>0$,

$$\sum_{n=1}^{\infty} e^{-\varepsilon n/(q(n))^2}<\infty$$

in order to achieve the conclusion in (3.1).

THEOREM 3.2. *Under the conditions of Theorem* 3.1 *it follows that,*

$$\int [\hat{f}(x)-f(x)]^2 dx \to 0 \ w.p.\ 1 \qquad as\ n\to\infty . \tag{3.8}$$

PROOF. Note that

$$\int [\hat{f}(x)-f(x)]^2 dx \leqq \sup_x |\hat{f}(x)-f(x)| \int |\hat{f}(x)-f(x)|dx$$

thus it follows from Theorem 3.1 that we need only to show that

$$\int |\hat{f}(x)-f(x)|dx \to \text{w.p. } 1 \qquad \text{as } n\to\infty .$$

Now,

$$\begin{aligned}\int |\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|dx &\leqq \int \left| n^{-1}\sum_{i=1}^{n} \Phi_n(x, X_i)-\mathrm{E}\,\Phi_n(x, X_1)\right| dx \\ &\leqq \sum_{j=0}^{q(n)} |\hat{\theta}_j-\theta_j| \leqq (q(n)+1)\sum_{i=1}^{n} Y_{in} ,\end{aligned} \tag{3.9}$$

where

$$Y_{in}=(q(n)+1)^{-1}\sum_{j=0}^{q(n)} |\phi_j(X_i)-\mathrm{E}\,\phi_j(X_i)| , \qquad i=1,\cdots, n .$$

Set $\mu_{1n}=\mathrm{E}\,Y_{1n}$. Then for any $\varepsilon_n\geqq 0$,

$$\begin{aligned}&\mathrm{P}\left[\int |\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|dx \geqq \varepsilon_n\right] \\ &\quad\leqq \mathrm{P}\left[n^{-1}\sum_{i=1}^{n}(Y_{in}-\mu_{1n}) \geqq n\varepsilon_n/C(q(n)+1)-\mu_{1n}\right] .\end{aligned} \tag{3.10}$$

Thus as in Theorem 3.1 we conclude that

$$\int |\hat{f}(x) - \mathrm{E}\,\hat{f}(x)|dx \to 0 \text{ w.p. } 1 \qquad \text{as } n \to \infty .$$

Next,

$$\int |\mathrm{E}\,\hat{f}(x) - f(x)|dx \leqq \sum_{j=q(n)+1}^{\infty} \theta_j \int |\phi_j(x)|dx , \tag{3.11}$$

where the last upper bound converges to 0 as $n \to \infty$ since $\phi_j(x)$ is absolutely integrable. The theorem is now proved in view of the inequality

$$\int |\hat{f}(x) - f(x)|dx \leqq \int |\hat{f}(x) - \mathrm{E}\,\hat{f}(x)|dx + \int |\mathrm{E}\,\hat{f}(x) - f(x)|dx .$$

4. Application to density functional estimation

In this section we are interested in estimating the functional $\Delta(f) = \int f^2(x)dx$. A recent work Schuler and Wolff [8] suggests the estimator $\tilde{\Delta}(f) = \int \hat{f}^2(x)dx$, where $\hat{f}(x)$ is the orthogonal series estimate of $f(x)$ given by (1.3). Alternatively one might suggest to estimate $\Delta(f)$ by $\hat{\Delta}(f) = \int \hat{f}(x)dF_n(x)$ where $F_n(x)$ is the empirical distribution function. This second estimator has the advantage of simplicity and it is asymptotically equivalent to $\hat{\Delta}(f)$. It is possible to show that $\mathrm{E}|\hat{\Delta}(f) - \Delta(f)|^2 \to 0$ as $n \to \infty$ exactly as in Schuler and Wolff [8]. Using Theorem 3.1 we shall prove that both $\hat{\Delta}(f)$ and $\tilde{\Delta}(f)$ are strongly consistent.

THEOREM 4.1. *Let the condition of Theorem* 3.1 *be satisfied. Then*

$$|\hat{\Delta}(f) - \Delta(f)| \to 0 \ \textit{w.p. } 1 \qquad \textit{as } n \to \infty , \tag{4.1}$$

and

$$|\tilde{\Delta}(f) - \Delta(f)| \to 0 \ \textit{w.p. } 1 \qquad \textit{as } n \to \infty . \tag{4.2}$$

PROOF. Let us prove (4.1) first.

$$\begin{aligned} |\hat{\Delta}(f) - \Delta(f)| &\leqq \int [\hat{f}(x) - f(x)]^2 dx + 2 \int |\hat{f}(x) - f(x)| f(x)dx \\ &\leqq \int [\hat{f}(x) - f(x)]^2 dx + 2 \sup_x |\hat{f}(x) - f(x)| . \end{aligned} \tag{4.3}$$

The first term of the last upper bound of (4.3) converges to 0 w.p. 1 as $n \to \infty$ from Theorem 3.2, while the second term converges to 0 w.p. 1 as $n \to \infty$ from Theorem 3.1. Next, let us prove (4.2),

$$
\begin{aligned}
(4.4)\qquad |\tilde{\Delta}(f)-\Delta(f)| &\leqq \int |\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|dF_n(x) \\
&\quad + \left|\int \mathrm{E}\,\hat{f}(x)dF_n(x)-\int f(x)dF_n(x)\right| \\
&\quad + \left|\int f(x)dF_n(x)-\int f(x)dF(x)\right| \\
&= J_1+J_2+J_3\,, \qquad \text{say}\,.
\end{aligned}
$$

By the strong law of large numbers, $J_3 \to 0$ w.p. 1 as $n \to \infty$. While

$$
(4.5)\qquad J_2 \leqq \int |\mathrm{E}\,\hat{f}(x)-f(x)|dF_n(x) \leqq \sup_x |\mathrm{E}\,\hat{f}(x)-f(x)|\,,
$$

since $\int dF_n(x)=1$, and we have $J_2 \to 0$ w.p. 1 as $n \to \infty$. Finally,

$$
(4.6)\qquad J_1 \leqq \sup_x |\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\,,
$$

which converges to 0 w.p. 1 as $n \to \infty$ by Theorem 3.1. This complete the proof.

Remark. Using the remark following Lemma 2.1 we have that in the i.i.d. case $\hat{\Delta}(f)$ and $\tilde{\Delta}(f)$ both strongly consistent estimates of $\Delta(f)$ if for any $\varepsilon>0$, $\sum_{n=1}^{\infty} \exp(-\varepsilon n/[q(n)]^2)<\infty$.

Acknowledgment

The author would like to thank the referee for some useful comments which resulted in improving the readability of the paper.

McMaster University

References

[1] Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.
[2] Cencov, N. N. (1962). Estimation of an unknown density function from observations, *Dokl. Akad. Nauk, SSSR*, **147**, 45-48 (in Russian).
[3] Dvoretzky, A. (1972). Asymptotic normality for sums of dependent random variables, *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*, **2**, 513-535.
[4] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Ass.*, **58**, 13-30.
[5] Ibragimov, I. A. and Linnik, Yu, V. (1971). *Independent and Stationary Sequences of Random Variables*, Walters-Noordhoff, the Netherlands.
[6] Krornmal, R. and Tarter, M. (1968). The estimation of probability densities and cumulative by Fourier series methods, *J. Amer. Statist. Ass.*, **63**, 925-952.
[7] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition, *Proc. Nat. Acad. Sci. USA*, **42**, 43-47.

[8] Schuler, L. and Wolff, H. (1976). Zur Schatzung eines dichtefunktionals, *Metrika*, **23**, 149-153.

[9] Schwartz, S. C. (1967). Estimation of a probability density by an orthogonal series, *Ann. Math. Statist.*, **38**, 1262-1265.

[10] Watson, G. S. (1969). Density estimation by orthogonal series, *Ann. Math. Statist.*, **40**, 1496-1498.