36

PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 9, Fasc. 2 (1988), p. 1–11

HAAR SYSTEM AND NONPARAMETRIC DENSITY ESTIMATION IN SEVERAL VARIABLES

BY

Z. CIESIELSKI (SOPOT)

Abstract. Partial sums of the Fourier-Haar expansion in several variables are used to esstimate on cubes a probability density satisfying some Lipschitz conditions.

1. Introduction. We restrict our attention to function spaces and probability densities on the unit cube $Q = I^d \subset R^d$, $I = \langle 0, 1\rangle$, $d = 1, 2, \ldots$ For given $m = 0, 1, \ldots$ denote by Q_m a family of dyadic cubes in R^d such that

(i) $$Q = \bigcup \{J: J \in Q_m\},$$

(ii) $$J \cap J' = \emptyset \quad \text{for } J \neq J',$$

(iii) $$|J| = 2^{-dm},$$

where $|J|$ is the d-dimensional volume of J. Now, for fixed m define the P_m: $L^1(Q) \to L^1(Q)$ by

$$P_m f(x) = \frac{1}{|J|}\int_J f(y)\,dy \quad \text{for } x \in J,\ J \in Q_m. \tag{1.1}$$

The function $P_m f$ can be also viewed as a partial sum of the Fourier-Haar expansion of f or else $\{P_m f, m = 0, 1, \ldots\}$ can be treated as a martingale (see. e.g. [5]). For the kernel corresponding to the operator (1.1) we have

$$P_m(x, y) = 2^{dm} \sum_{J \in Q_m} \chi_J(x)\chi_J(y), \qquad x, y \in I^d, \tag{1.2}$$

where χ_J is the indicator of J. Clearly,

$$P_m f(x) = \int_Q P_m(x, y) f(y)\,dy. \tag{1.3}$$

Since P_m is symmetric, nonnegative and $P_m 1 = 1$, it follows that

P_m: $D(Q) \to D(Q)$, where $D(Q)$ is the set of all probability densities concentrated on Q, i.e.

$$D(Q) = \{f \in L^1(Q): f \geqslant 0 \text{ on } Q, \int_Q f = 1\}.$$

It is also clear from (1.2) that $P_m(x, \cdot) \in D(Q)$ for fixed $x \in Q$.

We assume that we are given a probability space $(\Omega, \mathscr{F}, \Pr)$ and a simple sample of size n, i.e. a sequence $X_1, X_2, \ldots$ of i.i.d. random vectors with values in Q and such that their common distribution has a density $f \in D(Q)$. The standard way of producing estimators for f is given by formula

$$f_{m,n}(x) = \frac{1}{n} \sum_{j=1}^{n} P_m(x, X_j), \tag{1.5}$$

which can be written in the form

$$f_{m,n}(x) = \sum_{J \in Q_m} n(J)\, h_J(x) \tag{1.6}$$

with

$$n(J) = \frac{1}{n} |\{j: X_j \in J\}|, \qquad h_J(x) = \frac{1}{|J|} \chi_J(x).$$

Thus, the diagram of $f_{m,n}$: $Q \to R$ is simply the *histogram*. Our aim is to investigate the rate of convergence of $f_{m,n}$ to f as m and n go to infinity and f is a Lipschitz class. For those classes the optimal relation between m and n will be described. The first results in this direction we find in Glivenko's book [7] (see also [10]).

2. Preliminaries. We are going to discuss probability densities from $D(Q) \cap C(Q)$ and from $D(Q) \cap L^p(Q)$ with $1 \leqslant p < \infty$. To this end we need some properties of the operator P_m. The most elementary are the following:

$$P_m \geqslant 0, \tag{2.1}$$

$$P_m^2 = P_m, \tag{2.2}$$

$$P_m 1 = 1, \tag{2.3}$$

$$\|P_m f\|_p \leqslant \|f\|_p \quad \text{for } 1 \leqslant p \leqslant \infty, f \in L^p(Q), \tag{2.4}$$

where

$$\|f\|_p = \|f\|_{L^p(Q)} = \Big(\int_Q |f|^p\Big)^{1/p}, \quad \|f\|_\infty = \|f\|_{L^\infty(Q)} = \operatorname{ess\,sup}\{|f(x)|: x \in Q\}.$$

The modulus of smoothness of $f \in L^p(Q)$ is defined as

$$\omega_p(f; \delta) = \sup_{|h|_\infty < \delta} \Big(\int_{Q(h)} |f(x+h) - f(x)|^p\, dx\Big)^{1/p},$$

where $Q(h) = \{x \in Q\colon x+h \in Q\}$ and, for $f \in C(Q)$,

$$\omega_p(f;\delta) = \sup_{|x-y|_\infty < \delta; x,y \in Q} |f(x)-f(y)|,$$

where $|x|_\infty = \max(|x_1|, \ldots, |x_d|)$.

PROPOSITION 2.5. *For $f \in C(Q)$ we have*

$$\|f-P_m f\| \leqslant \omega_\infty\left(f; \frac{1}{2^m}\right). \tag{2.6}$$

Conversely, let for some nondecreasing $\omega\colon R_+ \to R_+$

$$\|f-P_m f\|_\infty \leqslant \omega\left(\frac{1}{2^m}\right) \quad \textit{for } m = 0, 1, \ldots \tag{2.7}$$

Then

$$\omega_\infty(f;\delta) \leqslant 4d\omega(2\delta) \quad \textit{for } \delta > 0. \tag{2.8}$$

Proof. Inequality (2.6) is a simple consequence of (1.1). The converse can be proved as follows. If for some $J \in \mathcal{Q}_m$ the points x', x'' are in J, then $P_m f(x') = P_m f(x'')$ and, by (2.7),

$$|f(x')-f(x'')| \leqslant |f(x')-P_m f(x')| + |f(x'')-P_m f(x'')| \leqslant 2\omega\left(\frac{1}{2^m}\right). \tag{2.9}$$

Since f is continuous, it follows that (2.9) holds for $x', x'' \in \bar{J}$. Let now $x', x'' \in Q$ be arbitrary two different points and let m be such that

$$\frac{1}{2^m} \geqslant |x'-x''|_\infty > \frac{1}{2^{m+1}}. \tag{2.10}$$

Since

$$x''-x' = \sum_{j=1}^{d} (y^{(j)} - y^{(j-1)}), \quad \text{where } y^{(j)} = \sum_{k=1}^{j} (x_k''-x_k') e_k,$$

with e_k being the k-th unit vector in R^d, we find by (2.10) that $y^{(j)}$ and $y^{(j-1)}$ belong to two neighbouring cubes from $\mathcal{Q}_m$ and therefore, by (2.9),

$$|f(x')-f(x'')| \leqslant 4d\omega(2|x'-x''|_\infty).$$

The converse part of Proposition 2.5 for $d = 1$ was proved in [2].

COROLLARY 2.11. *Let $0 < \alpha \leqslant 1$ and $f \in C(Q)$ be given. Then the following conditions are equivalent*:

$$\|f-P_m f\|_\infty = O\left(\frac{1}{2^{\alpha m}}\right) \quad \textit{as } m \to \infty, \tag{2.12}$$

$$\omega_\infty(f;\delta) = O(\delta^\alpha) \quad \textit{as } \delta \to 0_+. \tag{2.13}$$

The L^p-case is little more complicated. We have the following direct result:

PROPOSITION 2.14. *Let* $1 \leqslant p < \infty$ *and let* $f \in L^p(Q)$. *Then*

$$\|f - P_m f\|_p \leqslant 2^{d/p}\, \omega_p\left(f; \frac{1}{2^m}\right). \tag{2.15}$$

Proof. For $J \in Q_m$ we have

$$\int_J \left| f(x) - \frac{1}{|J|} \int_J f(y)\,dy \right|^p dx \leqslant \frac{1}{|J|} \iint_{J^2} |f(x) - f(y)|^p\,dx\,dy$$

$$= \frac{1}{|J|} \int_{2^m|h|_\infty \leqslant 1} dh \int_{J(h)} |\Delta_h f(y)|^p\,dy, \qquad \text{where } J(h) = \{x \in J\colon x + h \in J\}.$$

It follows that $J(h) \subset J \cap Q(h)$ and, therefore,

$$\|f - P_m f\|_p^p \leqslant 2^{dm} \int_{2^m|h|_\infty \leqslant 1} dh \int_{Q(h)} |\Delta_h f(y)|^p\,dy \leqslant 2^d \left(\omega_p\left(f; \frac{1}{2^m}\right)\right)^p.$$

The converse result depends on the following Bernstein type inequality (in case $d = 1$, see [3, 4], and for $d > 1$, [5]).

PROPOSITION 2.16. *Define*

$$S_m(Q) = \operatorname{span}\{\chi_J\colon J \in Q_m\}.$$

Then, for $1 \leqslant p < \infty$ *and for* $f \in S_m(Q)$, *we have*

$$\|\Delta_h f\|_{L^p(Q(h))} \leqslant 2d \cdot 3^{d/p} (2^m |h|_\infty)^{1/p} \|f\|_{L^p(Q)} \quad \textit{for } |h|_\infty \leqslant \frac{1}{2^m}. \tag{2.17}$$

Proof. Let $e_1, \ldots, e_d$ be the basic unit vectors in R^d and, for $h = (h_1, \ldots, h_d)$, let $h(j) = h_1 e_1 + \ldots + h_j e_j$. Since

$$\Delta_h f(x) = \sum_{j=1}^{d} \Delta_{h_j e_j} f\big(x + h(j-1)\big), \qquad h(0) = 0,$$

we obtain, for $J \in Q_m$,

$$\int_{J \cap Q(h)} |\Delta_h f|^p \leqslant d^{p-1} \sum_{j=1}^{d} \int_{J \cap Q(h)} |\Delta_{h_j e_j} f\big(x + h(j-1)\big)|^p\,dx$$

$$= d^{p-1} \sum_{j=1}^{d} \int_{J \cap Q(h)} \big|f\big(x + h(j)\big) - f\big(x + h(j-1)\big)\big|^p\,dx.$$

Now, $f\big(x + h(j)\big) = f\big(x + h(j-1)\big)$ for $x \in \big(J - h(j)\big) \cap \big(J - h(j-1)\big)$ and, therefore,

$$\int_{J \cap Q(h)} \big|f\big(x + h(j)\big) - f\big(x + h(j-1)\big)\big|^p\,dx$$

$$= \int_E \big|f\big(x + h(j)\big) - f\big(x + h(j-1)\big)\big|^p\,dx \leqslant 2^{p-1}\Big(\int_{E_{j-1}} |f|^p + \int_{E_j} |f|^p\Big),$$

where

$$E = J \cap Q(h) \backslash (J - h(j)) \cap (J - h(j-1))$$

and

$$E_j = E + h(j) = (J \cap Q(h) + h(j)) \backslash J \cap (J + h_j e_j).$$

Let now $J^* = \bigcup \{J_\varepsilon \colon \varepsilon = (\varepsilon_1, \ldots, \varepsilon_d), \varepsilon_j = 0, 1, -1\} \cap Q$, where $J_\varepsilon = J + \varepsilon/2^m$. It should be clear that $E_j = J^* \cap E_j = \bigcup_\varepsilon J_\varepsilon \cap E_j$. Now $J_\varepsilon = J$ for $\varepsilon = (0, \ldots, 0)$ and then

$$|J \cap E_j| \leqslant |J| - |J \cap (J + |h|_\infty e_j)| = |J|\, 2^m |h|_\infty .$$

For $\varepsilon \neq 0$, $|J_\varepsilon \cap J| = 0$ and

$$|J_\varepsilon \cap E_j| = |(J \cap Q(h) + h(j)) \cap J_\varepsilon|$$
$$\leqslant |(J + h(j)) \cap J_\varepsilon| \leqslant \frac{|h|_\infty}{2^{(d-1)m}} = |J|\,(2^m |h|_\infty),$$

therefore, for $J_\varepsilon \subset Q$,

$$\int_{E_j \cap J_\varepsilon} |f|^p = \frac{|E_j \cap J_\varepsilon|}{|J|} \int_{J_\varepsilon} |f|^p \leqslant (|h|_\infty 2^m) \int_{J_\varepsilon} |f|^p,$$

whence we infer that

$$\int_{Q(h)} |\Delta_h f|^p \leqslant (2d)^{p-1} |h|_\infty \sum_{j=1}^{d} \sum_{J_\varepsilon \subset Q} \Big(\int_{E_j \cap J_\varepsilon} |f|^p + \int_{E_{j-1} \cap J_\varepsilon} |f|^p \Big)$$
$$\leqslant (2d)^p 3^d |h|_\infty 2^m \int_Q |f|^p .$$

We are now in position, using a standard method from approximation theory, to prove the main converse result.

THEOREM 2.18. *Let* $1 \leqslant p < \infty$ *and let* $f \in L^p(Q)$. *Then*

$$\omega_p\Big(f; \frac{1}{2^m}\Big) \leqslant \frac{6d \cdot 3^{d/p}}{2^{m/p}} \sum_{i=0}^{m} 2^{i/p} \|f - P_i f\|_p . \tag{2.19}$$

Proof. We have

$$f = P_1 f + \sum_{j=1}^{m} f_j + (f - P_m f)$$

with $f_j = P_j f - P_{j-1} f$, whence, for $|h|_\infty \leqslant 1/2^m$,

$$\Delta_h f = \sum_{j=1}^{m} \Delta_h f_j + \Delta_h (f - P_m f),$$

$$\|\Delta_h f\|_{L^p(Q(h))} \leqslant \sum_{j=1}^{m} \|\Delta f_j\|_{L^p(Q(h))} + 2 \|f - P_m f\|_p .$$

Now, (2.17) gives

$$\|\Delta_h f_j\|_{L^p(Q(h))} \leqslant 2d\cdot 3^{d/p}(|h|_\infty 2^j)^{1/p}\|f_j\|_p$$
$$\leqslant 2d\cdot 3^{d/p}(|h|_\infty 2^j)^{1/p})\|f-P_j f\|_p+\|f-P_{j-1}f\|_p).$$

Combining these inequalities, we get (2.19).

COROLLARY 2.20. *Let* $f\in L^p(Q)$ *and let* α *and* p *be such that* $0<\alpha<1/p\leqslant 1$. *Then the following conditions are equivalent*:

(i) $$\omega_p(f;\delta)=O(\delta^\alpha)\quad \text{as } \delta\to 0_+,$$

(ii) $$\|f-P_m f\|_p = O\left(\frac{1}{2^{\alpha m}}\right)\quad \text{as } m\to\infty.$$

This result in the 1-dimensional case we find in [8] and in [4]. It should be also mentioned here that Properties (2.2) and (2.4) imply

(2.21) $$E_{m,p}(f)\leqslant \|f-P_m f\|_p\leqslant 2E_{m,p}(f)\quad \text{for } 1\leqslant p\leqslant\infty,$$

where $E_{m,p}(f)=\inf\{\|f-g\|_p\colon g\in S_m(Q)\}$.

3. Estimation of continuous densities. As in Introduction, we are given a sequence $(X_1, X_2, \ldots)$ of i.i.d. random vectors with values in Q and with the common density $f\in C(Q)\cap D(Q)$. The random function $f_{m,n}$ is defines as in (1.5). It will be shown that, for suitable dependence of n on m, the function $f_{m,n}$ is a good estimator for f. In what follows it is assumed that the sample size n is a dyadic natural. For given positive β the dependence of m on n is defined by

(3.1) $$n=2^\nu \quad\text{and}\quad m=[\beta\nu/d],$$

where ν is natural and $[x]$ is the integer part of x. In this particular situation the $f_{m,n}$ is denoted by $f_{\nu,\beta}$. It is important that β is asymptotically $\log N/\log n$ for large ν, N being the number of elements in Q_m and n the size of the sample. Our aim is, given properties of f, to determine the best β and then to compute N.

The main tool in the following discussion is the Bernstein inequality (cf. [9], p. 19);

LEMMA 3.2. *Let* Y_j $(j=1, 2, \ldots, n)$ *be independent random variables such that* $\Pr\{Y_k=1\}=y$, $\Pr\{Y_k=0\}=z$, $y+z=1$. *Then*

(3.3) $$\Pr\{|\sum_{j=1}^{n}(Y_j-y)|\geqslant 2\omega(nyz)^{1/2}\}\leqslant 2e^{-\omega^2}\quad \textit{for } 0\leqslant\omega\leqslant\frac{3}{2}(nyz)^{1/2}.$$

The rate of convergence of $\|f-f_{\nu,\beta}\|_\infty$ to zero as $\nu\to\infty$ can be investigated with the help of inequalities ($m=[\beta\nu/d]$, $1\leqslant p\leqslant\infty$)

(3.4) $$\frac{1}{2}\|f-P_m f\|_p\leqslant E_{m,p}\leqslant\|f-f_{\nu,\beta}\|_p\leqslant\|f-P_m f\|_p+\|P_m f-f_{\nu,\beta}\|_p,$$

which hold with probability 1 by the definition of $E_{m,p}$ and by (2.21).

LEMMA 3.5. *Let $f \in C(Q) \cap D(Q)$ and let $k > 0$, $\lambda > 0$, $0 < \beta < \frac{1}{2}$. Then*

$$\Pr\left\{\left\|\frac{P_m f - f_{\nu,\beta}}{P_m f}\right\|_\infty > \lambda\right\} = O(\varepsilon + \varepsilon^k 2^{md}), \quad m = \left[\frac{\beta\nu}{d}\right], \tag{3.6}$$

where $\varepsilon = \lambda^{-1} \cdot 2^{md(1-1/2\beta)}$ and the big O is independent of λ.

Proof. Note that

$$\left\|\frac{P_m f - f_{\nu,\beta}}{P_m f}\right\|_\infty = \sup_{J \in Q_m} \frac{\left|\int_J f - n^{-1} \sum_{j=1}^{n} \chi_J(X_j)\right|}{\int_J f}. \tag{3.7}$$

Now, for $J \in Q_m$, we put $Y_j = \chi_J(X_j)$, $y = \int_J f$, $y+z = 1$, and then apply Lemma 3.2 to get (3.3) with $\omega = \lambda ny/2\sqrt{nyz} \leqslant \frac{3}{2}\sqrt{nyz}$, provided that $\lambda \leqslant 3z$. This condition holds in particular for λ and m satisfying

$$2^{dm} \geqslant \tfrac{3}{2}\|f\|_\infty, \quad \lambda \leqslant 1. \tag{3.8}$$

Now,

$$\omega^2 = \frac{\lambda^2 yn}{4z} \geqslant \lambda_1 \frac{1}{|J|} \int_J f, \quad \lambda_1 = \frac{\lambda^2 n}{4 \cdot 2^{md}},$$

and, therefore, by Jensen's inequality

$$\sum_{J \in Q_m} e^{-\omega^2} \leqslant \sum_{J \in Q_m} e^{(-\lambda_1/|J|)\int_J f} \leqslant 2^{md} \sum_{J \in Q_m} \int_J e^{-\lambda_1 f(x)}\, dx \tag{3.9}$$

$$= 2^{md} \int_Q e^{-\lambda_1 f(x)}\, dx = 2^{md} \int_0^\infty e^{-\lambda_1 s}\, dF_f(s),$$

where F_f is the distribution of f on $\langle 0, \infty)$ with respect to the Lebesgue measure on Q. Now,

$$\lambda_1 \geqslant \frac{\lambda^2}{4} 2^{md(1/\beta - 1)},$$

whence, for $\gamma > 0$,

$$N \int_0^\infty e^{-\lambda^2 N^{1/\beta - 1} s/4}\, dF_f(s) \leqslant N\Big(\int_0^{N^{-\gamma}\lambda^{-1}} + \int_{N^{-\gamma}\lambda^{-1}}^\infty\Big) e^{-\lambda^2 N^{1/\beta-1} s/4}\, dF_f(s) \tag{3.10}$$

$$\leqslant \frac{1}{\lambda} N^{1-\gamma} + N e^{-\lambda N^{1/\beta - 1 - \gamma}/4} \leqslant \frac{1}{\lambda} N^{1-\gamma} + N\left(\frac{\lambda}{4} N^{1/\beta - 1 - \gamma}\right)^{-k} \sup_{0<x<\infty} x^k e^{-x}$$

$$= O\left(\frac{1}{\lambda} N^{1-\gamma} + \frac{1}{\lambda^k} N^{1+k(1+\gamma-1/\beta)}\right),$$

where $N = 2^{md}$, k is any positive number and the O depends on β, d, and k only. Combining (3.7)–(3.10) with $\gamma = 1/2\beta$ we obtain (3.6).

PROPOSITION 3.11. *Let* $f \in C(Q) \cap D(Q)$ *and let* $0 < \beta < 1/2$. *Then*

$$\Pr\{\|f - f_{v,\beta}\|_\infty = o(1) \text{ as } v \to \infty\} = 1. \tag{3.12}$$

Proof. It follows by (3.6) that taking $k > 0$ such that $1/2\beta - 1 > 1/k$, we obtain with probability 1 for large m

$$\bigwedge_{x \in Q} \quad |P_m f(x) - f_{v,\beta}(x)| \leqslant \lambda |P_m f(x)|,$$

whence $\|P_m f - f_{v,\beta}\|_\infty \leqslant \lambda \|f\|_\infty$, and therefore

$$\Pr\{\|P_m f - f_{v,\beta}\|_\infty = o(1) \text{ as } v \to \infty\} = 1.$$

On the other hand, according to (2.15), $\|f - P_m f\|_\infty = o(1)$ as $m \to \infty$. Thus, (3.4) implies (3.12).

THEOREM 3.13. *Let* $f \in C(Q) \cap D(Q)$. *Then for* $0 < \alpha \leqslant 1$, $0 < \beta < d/2(\alpha + d)$ *the following conditions are equivalent*:

(i) $$\omega_\infty(f; \delta) = O(\delta^\alpha) \quad \text{as } \delta \to 0_+,$$

(ii) $$\Pr\left\{\|f - f_{v,\beta}\|_\infty = O\left(\frac{1}{2^{m\alpha}}\right) \text{ as } m \to \infty\right\} = 1, \qquad m = \left[\frac{v\beta}{d}\right].$$

Proof. (i) $\Rightarrow$ (ii). According to Corollary 2.5 we have $\|f - P_m f\|_\infty = O(1/2^{m\alpha})$, and (3.6) with $\lambda = 1/2^{m\alpha}$ and k such that $(k-1)(1/2\beta - 1 - \alpha/d) \geqslant 1$ gives

$$\Pr\left\{\|P_m f - f_{v,\beta}\|_\infty > \frac{1}{2^{m\alpha}}\right\} = O\left(2^{md(\alpha/d + 1 - 1/2\beta)}\right).$$

Combining these inequalities with (3.4) we complete this part of the proof.

(ii) $\Rightarrow$ (i). Using (3.4) we find that $\|f - P_m\|_\infty = O(1/2^{\alpha m})$, whence by Proposition 2.5 the required result follows.

4. Estimation of densities in L^p. Like in the previous section we consider densities concentrated on the d-dimensional cube Q. It is also assumed that (3.1) is satisfied. The expectation of an r.v. Y with respect to the given probability space $(\Omega, \mathscr{F}, \Pr)$ is denoted by EY.

The following result from Lorentz and Berens [1] plays an important role in our considerations:

PROPOSITION 4.1. *Let* $g \in C(I)$, $I = \langle 0, 1 \rangle$. *Then, for* $x \in I$,

$$\left| g(x) - \sum_{j=0}^{n} g\left(\frac{j}{n}\right) \binom{n}{j} x^j (1-x)^{n-j} \right| \leqslant 3\omega_{2,\infty}\left(g; \frac{1}{2}\sqrt{\frac{x(1-x)}{n}}\right),$$

where

(4.2) $$\omega_{2,\omega}(g;\delta) = \sup_{\substack{x_1,x_2\in I\\ |x_1-x_2|\leqslant 2\delta}} \left| g\left(\frac{x_1+x_2}{2}\right) - \frac{g(x_1)+g(x_2)}{2}\right|, \quad 0<\delta\leqslant\frac{1}{2}.$$

The following elementary inequalities are well known.

PROPOSITION 4.3. *Let* $I = \langle -1, 1\rangle$, $R = (-\infty, \infty)$. *Then*

(i) $0 \leqslant |x+h|^p + |x-h|^p - 2|x|^p \leqslant 2|h|^p$ *for* $1\leqslant p\leqslant 2$, $x+h$, $x-h\in R$,

(ii) $0 \leqslant |x+h|^p + |x-h|^p - 2|x|^p \leqslant p(p-1)|h|^2$ *for* $p>2$, $x+h$, $x-h\in I$.

PROPOSITION 4.4. *Let* $\beta>0$, $1\leqslant p<\infty$ *and let* $f\in L^p(Q)\cap D(Q)$. *Then*

(4.5) $$\|f-P_m f\|_p \leqslant (E\|f-f_{\nu,\beta}\|_p^p)^{1/p} \leqslant \|f-P_m f\|_p + (E\|P_m f - f_{\nu,\beta}\|_p^p)^{1/p}.$$

Proof. Since $Ef_{\nu,\beta}(x) = P_m f(x)$, Jensen's inequality implies the first inequality in (4.5). The second one follows by the triangle inequality.

LEMMA 4.6. *Let* $1\leqslant p<\infty$, $p^{-1}+q^{-1}=1$, $\beta>0$ *and let* $f\in L^p(Q)\cap D(Q)$. *Then, under* (3.1),

(4.7) $$(\mathrm{E}\|P_m f - f_{\nu,\beta}\|_p^p)^{1/p} \leqslant C\cdot 2^{-md\gamma} \quad \text{for } \nu\to\infty,$$

where

$$\gamma = \frac{1}{\beta}\frac{1}{2\vee p} - \frac{1}{2\wedge q}$$

($a\wedge b = \min(a,b)$, $a\vee b = \max(a,b)$), *and* C *depends on* p *only.*

Proof. Notice that with $N = 2^{md}$ we have

(4.8) $$\mathrm{E}\|P_m f - f_{m,\beta}\|_p^p = N^{p-1}\sum_{J\in Q_m} \mathrm{E}\left|\frac{1}{n}\sum_{j=1}^{n}\big(Y_j(J)-y(J)\big)\right|^p,$$

where $Y_j(J) = \chi_J(X_j)$, $y(J) = \Pr\{Y_j(J)=1\} = \mathrm{E}Y_j(J) = \int_J f$. Applying Propositions 4.1 and 4.3 to $g(x) = |x-y(J)|^p$, we obtain

(4.9) $$\mathrm{E}\left|\frac{1}{n}\sum_{j=1}^{n}\big(Y_j(J)-y(J)\big)\right|^p \leqslant C_1\left(\frac{y(J)\big(1-y(J)\big)}{n}\right)^{(2\wedge p)/2}.$$

The combination of (4.9) and (4.8) gives

(4.10) $$\mathrm{E}\|P_m f - f_{m,\beta}\|_p^p \leqslant C_2 N^{p-1} n^{-(2\wedge p)/2} \sum_{J\in Q_m}\big(y(J)(1-y(J))\big)^{(2\wedge p)/2}.$$

Now, $1-y(J)\leqslant 1$ and in addition, by concavity, for $1\leqslant p\leqslant 2$ we have

$$\sum_{J\in Q_m} y(J)\big(1-y(L)\big)^{(2\wedge p)/2} \leqslant \sum_{J\in Q_m} y(J)^{p/2} \leqslant N^{1-p/2}\Big(\sum_{J\in Q_m} y(J)\Big)^{p/2} = N^{1-p/2},$$

and, for $p > 2$,

$$\sum_{J \in \mathcal{Q}_m} \big(y(J)(1-y(J))\big)^{(2 \wedge p)/2} \leqslant \sum_{J \in \mathcal{Q}_m} y(J) = 1.$$

Both these inequalities and (4.10) give

$$\mathrm{E}\,\|P_m f - f_{\nu,\beta}\|_p^p \leqslant C_3^p N^{-p\gamma}. \tag{4.11}$$

Now, we are in position to state our main theorem for the L^p-case, namely

THEOREM 4.12. *Assume* (3.1) *and*

$$0 < \alpha \leqslant 1, \quad 1 \leqslant p < \infty, \quad 0 < \beta < \frac{d}{(2 \vee p) + ((2 \vee p) - 1)d}. \tag{4.13}$$

Then, for $f \in L^p(Q) \cap D(Q)$, *the following conditions are equivalent*:

(i) $$\|f - P_k f\|_p = O\left(\frac{1}{2^{k\alpha}}\right) \quad \text{as } k \to \infty,$$

(ii) $$(\mathrm{E}\,\|f - f_{\nu,\beta}\|_p^p)^{1/p} = O\left(\frac{1}{2^{\nu\alpha\beta/d}}\right) \quad \text{as } \nu \to \infty.$$

Moreover, next condition implies (i):

(iii) $$\omega_p(f;\delta) = O(\delta^\alpha) \quad \text{as } \delta \to 0_+.$$

If, in addition to (4.13), $0 < \alpha < 1/p$, *then also* (i) *implies* (iii).

Proof. In view of Corollary 2.7 it is sufficient to show that (α) the model is regular, and (β) each limit of UBE's is admissible.

$$\frac{1}{2}\|f - Pf_m\|_p \leqslant (\mathrm{E}\,\|f - f_{\nu,\beta}\|_p^p)^{1/p} \leqslant 2\big(\|f - P_m f\|_p + (\mathrm{E}\,\|P_m f - f_{\nu,\beta}\|_p^p)^{1/p}\big). \tag{4.14}$$

This, (3.1), (i) and Lemma 4.6 imply

$$(\mathrm{E}\,\|f - f_{\nu,\beta}\|_p^p)^{1/p} \leqslant O\left(\frac{1}{2^{m\alpha}} + \frac{1}{2^{md\gamma}}\right),$$

where $\gamma = 1/(2 \vee p)/\beta - 1/(2 \wedge q)$. From this (ii) follows.

(ii) $\Rightarrow$ (i). It follows by (4.14) that (i) is satisfied for $k = m = [\beta\nu/d]$. However, by (4.13), $\beta/d < 1$ and therefore each k is of the form $[\beta\nu/d]$.

(iii) $\Rightarrow$ (i). This implication holds true by Proposition 2.14. Its converse in case $0 < \alpha < 1/p$ follows by Corollary 2.20.

COROLLARY 4.15. *Let* $f \in L^p(Q) \cap D(Q)$ *for some* p $(1 \leqslant p \leqslant \infty)$ *and let* (iii) *hold for some* α $(0 < \alpha \leqslant 1)$. *Then, for each* β *satisfying* (4.13), *we have*

$$\Pr\{\|f - f_{\nu,\beta}\|_p \to 0 \text{ as } \nu \to \infty\} = 1.$$

COROLLARY 4.16. *For given α, p and βsatisfying* (4.13) *the best choice for β with respect to* (ii) *is*

$$\beta = \frac{d}{(2 \vee p)\alpha + ((2 \vee p) - 1)d}.$$

Examples. 1. Let α, β and p be as in (4.13). Let $f \in W^1_{p'}(Q)$ for some p' satisfying the inequalities $1 \leqslant p' \leqslant p < \infty$ and $d(1/p' - 1/p) < 1$. Then $\omega_p(f; \delta) = O(\delta^\alpha)$ with $\alpha = 1 - d(1/p' - 1/p)$, and the β can be easily computed. This is actually an embedding theorem which can be derived for instance from [6].

2. Let $d = 1$ and $1 \leqslant p < 2$. Then the density for the arcsin law is given by the formula

$$f(x) = \frac{1}{\pi} \frac{1}{\sqrt{x(1-x)}}, \qquad x \in Q = \langle 0, 1 \rangle.$$

One checks that $f \in L^p, f \notin L^2$ and $\omega_p(f; \delta) = O(\delta^{1/p - 1/2})$. In this case $\alpha = 1/p - 1/2$ and $\beta = p/2$.

REFERENCES

[1] H. Berens and G. Lorentz, *Inverse theorems for Bernstein polynomials*, Ind. Univ. M. J. 21 (1973), p. 693–708.

[2] Z. Ciesielski, *On Haar functions and on the Schauder basis of the space* $C_{\langle 0,1 \rangle}$, Bull. Acad. Polon. Sci. 7 (1959), p. 227–232.

[3] — *Properties of the orthonormal Franklin system, II*, Studia Math. 27 (1966), p. 290–323.

[4] — *Constructive function theory and spline systems*, ibidem 53 (1975), p. 277–302.

[5] — *Haar orthogonal functions in analysis and probability*, Coll. Math. Soc. J. Bolyai 49, Alfred Haar Memorial Conference, Budapest 1985, p. 25–55.

[6] Z. Ciesielski and T. Figiel, *Spline bases in classical function spaces on compact C^∞ manifolds*, Studia Math. 75 (1983); Part I, p. 1–58, Part II, p. 96–136.

[7] V. I. Glivenko, *Course in probability theory*, Moscow — Leningrad 1939 (in Russian).

[8] B. I. Golubov, *Haar series* in : *Itogi Nauki*, Moskva 1971, p. 104–146 (in Russian).

[9] G. G. Lorentz, *Bernstein polynomials*, Toronto 1953.

[10] E. A. Nadaraja, *Nonparametric estimation of density and of regression curve*, Tibilisi 1983 (in Russian).

Instytut Matematyczny PAN
ul. Abrahama 18
81-825 Sopot
Poland

Received on 28. 3. 1987