Ann. Inst. Statist. Math.
34 (1982), Part A, 457-466

L_p-CONSISTENCY OF MULTIVARIATE DENSITY ESTIMATES

IBRAHIM A. AHMAD

(Received May 15, 1979; revised March 8, 1982)

Abstract

L_p notion of the weak, mean, and strong consistency of the kernel method of multivariate density estimation is proposed and studied. The results expand, unify, or generalize most known results in the literature. Rates of convergence in mean and strong L_p-consistencies are presented.

1. Introduction

Nonparametric density estimation has received a good deal of attention recently. Various approaches are suggested and studied for density estimation from a nonparametric viewpoint; we refer the reader to a paper by Fryer [3] for a recent review of the subject.

One such technique for estimating a probability density function is the so-called "kernel method" originated by Rosenblatt [10] and developed rigorously by Parzen [9] and Nadaraya [8], among many, for the univariate case and extended to the multidimensional case by Cacoullos [1], Van Ryzin [13], and Rüschendorf [11], among others.

Let $X_1, \cdots, X_n$ be a random sample from an m-dimensional distribution ($m \geq 1$) with probability density function (pdf) $f(x)$ and distribution function (df) $F(x)$. Let $k(u)$ be a known pdf defined on the m-dimensional Euclidean space R^m and satisfying the following conditions:

$$\sup_u k(u) < \infty \quad \text{and} \quad ||u|| k(u) \to 0 \quad \text{as } ||u|| \to \infty . \tag{1.1}$$

Also let $\{a_n\}$ be a sequence of real numbers such that $a_n \to 0$ as $n \to \infty$. Further conditions on k and a_n will be stated in the sequel. The kernel estimator of $f(x)$ is defined by:

AMS 1970 subject classification: Primary 62G05.
Key words and phrases: Density estimation, kernel method, L_p-consistency, rates of convergence.

$$(1.2)\qquad \hat{f}(x)=a_n^{-m}\int k[(x-u)/a_n]dF_n(u)=(na_n^m)^{-1}\sum_{i=1}^{n}k[(x-X_i)/a_n]\,,$$

where $k[(w-y)/a_n]=k[(w_1-y_1)/a_n, \cdots, (w_m-y_m)/a_n]$.

When discussing the consistency of the estimator (1.2) most authors were concerned with pointwise or L_∞ (uniform)-consistency, see, e.g., Parzen [9], Nadaraya [8], Cacoullos [1], Van Ryzin [13], and Rüschendorf [11] among others. Leadbetter [5] discussed the L_2-consistency by showing that under some conditions $\mathrm{E}\int[\hat{f}(x)-f(x)]^2dx$ converges to 0 as $n\to\infty$ when $m=1$. The purpose of the present investigation is to define a L_p-consistency in probability, in the mean, and strongly and discuss under what conditions (on a_n, k, and f) is $f(x)$ consistent, $1\leqq p\leqq\infty$. First let us give a definition.

DEFINITION 1.1. For any $1\leqq p<\infty$, let $\|\hat{f}-f\|_p=\left(\int|\hat{f}(x)-f(x)|^pdx\right)^{1/p}$ and for $p=\infty$, let $\|\hat{f}-f\|_\infty=\sup_x|\hat{f}(x)-f(x)|$. It is said that $\hat{f}$ is weakly L_p-consistent if $\|\hat{f}-f\|_p\to 0$ in probability as $n\to\infty$, that $\hat{f}$ is mean L_p-consistent if $\|\hat{f}-f\|_p\to 0$ in the mean as $n\to\infty$, and that $\hat{f}$ is strongly L_p-consistent if $\|\|\hat{f}-f\|_p\to 0$ with probability one as $n\to\infty$.

The purpose of this note is two-fold; first to discuss L_p-consistency of $\hat{f}$ for any p, $1\leqq p\leqq\infty$ and second to discuss the rates of convergence in L_p-consistency. We shall attempt to present the results using the most widely known conditions. Precisely, we shall establish mean L_p-consistency under conditions similar to those of Parzen [9] and Cacoullos [1] and establish strong L_p-consistency under conditions similar to those of Nadaraya [8] and Rüschendorf [11]. This is done in Section 2. Allowing slightly stronger conditions we establish in Section 3 the rates of convergence in mean and strong L_p-consistencies. Only mean and strong consistency will be discussed. Note that weak consistency follows directly from mean consistency. Throughout this paper, C (sometimes subscripted) denotes generic positive constants not necessarily the same.

2. L_p-consistency

THEOREM 1. (i) *If $f\in L_p$, $1\leqq p<\infty$, and if $na_n^m\to\infty$ as $n\to\infty$, then*

$$(2.1)\qquad \|\hat{f}-f\|_p\to 0 \qquad \text{in the mean as } n\to\infty\,.$$

(ii) *If $p=\infty$, if $na_n^{2m}\to\infty$, and if f is uniformly continuous,*

$$(2.2)\qquad \|\hat{f}-f\|_\infty\to 0 \qquad \text{in the mean as } n\to\infty\,.$$

PROOF. Note that (ii) is proved in Theorem 3.3 of Cacoullos [1]. Let us prove (i).

$$||\hat{f}-f||_p^p \leqq C_p\{||\hat{f}-\mathrm{E}\,\hat{f}||_p^p+||\mathrm{E}\,\hat{f}-f||_p^p\}\,. \tag{2.3}$$

Now, for $1\leqq p<\infty$

$$||\mathrm{E}\,\hat{f}-f||_p^p=\int|\mathrm{E}\,\hat{f}(x)-f(x)|^p dx \to 0 \qquad \text{as } n\to\infty\,, \tag{2.4}$$

since $|\mathrm{E}\,\hat{f}(x)-f(x)|^p\to 0$ as $n\to\infty$ provided x is a continuity point of f and that $na_n^m\to\infty$, see Cacoullos [1], and since $|\mathrm{E}\,\hat{f}(x)-f(x)|^p\leqq C_p[(\mathrm{E}\,\hat{f}(x))^p+(f(x))^p]$ which is integrable and converges at every continuity point x of f to $2C_p(f(x))^p$, an integrable function, thus by Lebesgue dominated convergence theorem, (2.4) follows. Next,

$$\mathrm{E}\,||\hat{f}-\mathrm{E}\,\hat{f}||_p\leqq \mathrm{E}^{1/p}\,||\hat{f}-\mathrm{E}\,\hat{f}||_p^p=\left\{\mathrm{E}\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|^p dx\right\}^{1/p}. \tag{2.5}$$

But

$$\begin{aligned}&\mathrm{E}\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|^p dx \\ &\leqq \mathrm{E}\,\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|^{p-1}\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|dx \\ &\leqq \mathrm{E}^{1/2}\,\{\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\}^{2p-2}\,\mathrm{E}^{1/2}\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|^2 dx\,.\end{aligned} \tag{2.6}$$

But it follows that for $0\leqq\nu<2$,

$$\mathrm{E}\,\{\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\}^\nu\leqq \mathrm{E}^{\nu/2}\,\{\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\}^2\,,$$

which converges to 0 as $n\to\infty$, see Theorem 3.3 of Cacollous [1]. This proof is an easy extension of the proof of (3.9) of Parzen [9]. Next, note that $\mathrm{E}\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|^2dx\to 0$ as $n\to\infty$ by an application of the Lebesgue dominated convergence theorem. Thus, it remains to show that for any $\nu\geqq 2$, $\mathrm{E}\,\{\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\}^\nu\to 0$ as $n\to\infty$. Let $\eta_n(t)=\frac{1}{n}\sum_{j=1}^n e^{it'X_j}$ and $\phi(t)=\int e^{-it'x}k(x)dx$. Then it follows that,

$$\hat{f}(x)=\left(\frac{1}{2\pi}\right)^m\int e^{-it'x}\phi(a_n\cdot t)\eta_n(t)dt\,. \tag{2.7}$$

Hence

$$\{\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\}^\nu\leqq\left(\frac{1}{2\pi}\right)^{\nu m}\left\{\frac{1}{n}\sum_{j=1}^n\int|\phi(a_n\cdot t)||e^{it'X_j}-\mathrm{E}\,e^{it'X_j}|dt\right\}^\nu \tag{2.8}$$

$$=\left(\frac{1}{2\pi}\right)^{\nu m}\left\{\frac{1}{n}\sum_{j=1}^{n} Y_{jn}\right\}^{\nu}, \quad \text{say.}$$

Note that $|Y_{jn}|\leqq 2\int|\phi(a_n\cdot t)|dt$, $j=1, \cdots, n$ and $Y_{1n}, \cdots, Y_{nn}$ are iid random variables. Thus by a theorem of Dharmadhikari, Fabian, and Jogdeo [2] we have that for any $\nu\geqq 2$ and all n,

$$\mathrm{E}\left\{\frac{1}{n}\sum_{j=1}^{n} Y_{jn}\right\}^{\nu}\leqq C_{m\nu}(na_n^m)^{-\nu/2}.$$

Hence

$$\mathrm{E}\,\{\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\}^{\nu}\leqq C_{m\nu}(na_n^m)^{-\nu/2}. \tag{2.9}$$

Then part (i) follows from (2.4), (2.6) and (2.9).

THEOREM 2. *Let k be a function of bounded variations and assume that for any $\varepsilon>0$, $\sum_{n=1}^{\infty}\exp\{-\varepsilon na_n^{2m}\}<\infty$.*

(i) *If $f\in L_p$, $1\leqq p<\infty$, then*

$$\|\hat{f}(x)-f(x)\|_p\to 0 \qquad \textit{with probability one as } n\to\infty. \tag{2.10}$$

(ii) *If $p=\infty$, f is uniformly continuous, then*

$$\|\hat{f}(x)-f(x)\|_\infty\to 0 \qquad \textit{with probability one as } n\to\infty. \tag{2.11}$$

PROOF. Again part (ii) is proved by Rüschendorf [10], Theorem 1(A). Let us prove (ii). Again notice that

$$\|\hat{f}(x)-f(x)\|_p^p\leqq C_p\{\|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)\|_p^p+\|\mathrm{E}\,\hat{f}(x)-f(x)\|_p^p\}.$$

The second term of the upper bound is shown in Theorem 1 to converge to 0 as $n\to\infty$. Next,

$$\begin{aligned}(2.12)\quad &\|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)\|_p^p\\ &=\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|^p dx\leqq\{\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\}^{p-1}\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|dx.\end{aligned}$$

But it is shown in Rüschendorf [11] that if k is a function of bounded variations and if for any $\varepsilon>0$, $\sum_{n=1}^{\infty}\exp\{-\varepsilon na_n^{2m}\}<\infty$, then $\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|\to 0$ with probability one as $n\to\infty$, hence we need only to show that $\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|dx\to 0$ with probability one as $n\to\infty$. Let $\mathscr{B}(0,\eta)$ denote the m-dimensional ball centered at the origin and of radius $\eta>0$. Thus we have

$$\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|dx \tag{2.13}$$

$$=\int_{\mathscr{B}(0,\eta)}|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|dx+\int_{\mathscr{B}^c(0,\eta)}|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|dx$$
$$\leqq \mathrm{V}\,(\mathscr{B}(0,\eta))\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|+\int_{\mathscr{B}^c(0,\eta)}|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|dx\,,$$

where $\mathrm{V}\,(\mathscr{B}(0,\eta))$ denote the volume of $\mathscr{B}(0,\eta)$. Let us deal with the second term in the rhs of (2.13).

$$\text{(2.14)}\qquad \int_{\mathscr{B}^c(0,\eta)}|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|dx$$
$$\leqq\int_{\mathscr{B}^c(0,\eta)}\hat{f}(x)dx+\int_{\mathscr{B}^c(0,\eta)}\mathrm{E}\,\hat{f}(x)dx$$
$$=-\int_{\mathscr{B}(0,\eta)}[\hat{f}(x)-\mathrm{E}\,\hat{f}(x)]dx+2\int_{\mathscr{B}^c(0,\eta)}\mathrm{E}\,\hat{f}(x)dx$$
$$\leqq\mathrm{V}\,(\mathscr{B}(0,\eta))\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|+2\int_{\mathscr{B}^c(0,\eta)}\mathrm{E}\,\hat{f}(x)dx\,,$$

since $\int\hat{f}(x)dx=1$ and $\int\mathrm{E}\,\hat{f}(x)dx=1$. But it is not difficult but tedious to show that for any $\delta>0$ we can choose η such that for n sufficiently large,

$$\text{(2.15)}\qquad \int_{\mathscr{B}^c(0,\eta)}\mathrm{E}\,\hat{f}(x)dx<\delta\,.$$

From (2.13)–(2.15) it follows that for sufficiently large η and large n we have

$$\text{(2.16)}\qquad ||\hat{f}-\mathrm{E}\,\hat{f}||_1\leqq\mathrm{V}\,(\mathscr{B}(0,\eta))||\hat{f}-\mathrm{E}\,\hat{f}||_\infty\,.$$

The desired conclusion now follows.

In the next remark we establish the necessity of the conditions of Theorem 2 above. We remark that for $m=1$, and $p=\infty$ the result appears in Schuster [12].

Remark 1. Assume that k is a function of bounded variation and that for any $\varepsilon>0$, $\sum_{n=1}^{\infty}\exp(-\varepsilon n a_n^{2m})<\infty$. A necessary condition that

$$\text{(2.17)}\qquad ||\hat{f}-g||_p\to 0\qquad\text{with probability one as } n\to\infty$$

for some measurable m-dimensional function g in L_p;

(i) When $1\leqq p<\infty$, then g is the L_p pdf of F.

(ii) When $p=\infty$, then g is the uniformly continuous pdf of F.

PROOF. (i) For $1\leqq p<\infty$ and using the C_p-inequality, see Loeve [6] p. 153, it follows that:

$$\text{(2.18)}\qquad ||\mathrm{E}\,\hat{f}-g||_p^p\leqq C_p\{||\hat{f}-\mathrm{E}\,\hat{f}||_p^p+||\hat{f}-g||_p^p\}\to 0\qquad\text{w.p. } 1\,,$$

as $n \to \infty$, by the argument of Theorem 2 and the assumption (2.17). Hence $\int |\mathrm{E}\,\hat{f}(x)-g(x)|^p dx \to 0$ as $n \to \infty$ and thus $\int (\mathrm{E}\,\hat{f}(x))^p dx \to \int g^p(x)dx$ as $n \to \infty$. Next, if $a'=(a_1, \cdots, a_m)$ and $b'=(b_1, \cdots, b_m)$, then $\int_a^b (\mathrm{E}\,\hat{f}(x))^p \cdot dx \to \int_a^b g^p(x)dx$ as $n \to \infty$. But from Fubini's theorem and the Lebesgue convergence theorem

$$\text{(2.19)} \qquad \int_a^b \mathrm{E}\,\hat{f}(x)dx \to F(b)-F(a_1, b_2, \cdots, b_m)+\cdots+(-1)^m F(a)\,.$$

Hence $g=f$ in the Lebesgue sense and by the fundamental theorem of calculus $f=g$ in the Rieman sense. To prove (ii) we proceed as follows:

$$\text{(2.20)} \qquad \|\mathrm{E}\,\hat{f}-g\|_\infty \leqq \|\hat{f}-\mathrm{E}\,\hat{f}\|_\infty + \|\hat{f}-g\|_\infty \to 0$$
$$\text{with probability one as } n \to \infty\,,$$

from Theorem 2 and the assumption of Remark 1. Thus F is continuous, since if not, let x_0 be a discontinuity point of F, i.e., $\mathrm{P}\,[X=x_0]>0$, then

$$\text{(2.21)} \qquad \sup_x \mathrm{E}\,\hat{f}(x) \geqq a_n^{-m}\,\mathrm{P}\,[X=x_0] \sup_x k[(x-x_0)/a_n]\,.$$

But also we have that

$$\text{(2.22)} \qquad \sup_x \mathrm{E}\,\hat{f}(x) \leqq a_n^{-m} M\,,$$

where $M=\sup_u k(u)$, this contradicts (2.20). Hence F is uniformly continuous, which implies that $\mathrm{E}\,\hat{f}(x)$ is uniformly continuous, but $g(x)$ is the uniform limit of $\mathrm{E}\,\hat{f}(x)$ then it is uniformly continuous, and hence $\int_a^b \mathrm{E}\,\hat{f}(x)dx \to \int_a^b g(x)dx$, but using Fubini's theorem and the Lebesgue convergence theorem (2.19) holds and thus $f=g$ in Lebesgue sence and by the fundamental theorem of calculus in the Rieman sense.

Remark 2. Another metric that may be used as a criterion for density estimate consistency is the so-called, Hellinger distance cf. Matusita [7], defined by

$$\text{(2.23)} \qquad \|\hat{f}^{1/2}-f^{1/2}\|_{H,p} = \left(\int |\hat{f}^{1/2}-f^{1/2}|^p dx\right)^{1/p} \qquad \text{for any } p \geqq 2\,.$$

The case $p=2$ is the usual definition of the Hellinger metric. Note that for any $p \geqq 2$

$$\|\hat{f}^{1/2}-f^{1/2}\|_{H,p}^p = \int |\hat{f}^{1/2}-f^{1/2}|^p dx \leqq \int |\hat{f}-f|^{p/2} dx = \|\hat{f}-f\|_{p/2}^{p/2}\,.$$

Hence $||\hat{f}^{1/2}-f^{1/2}||_{H,p}\leqq(||\hat{f}-f||_{p/2})^{1/2}$ and we conclude that Theorems 1 and 2 remain valid for the Hellinger metric.

3. Rates of convergence

We start this section by evaluating $\mathrm{E}\,||\hat{f}-f||_p^p$ for $1\leqq p\leqq\infty$. First let $1\leqq p<\infty$, and note that by the C_p-inequality,

$$\text{(3.1)}\qquad \mathrm{E}\int|\hat{f}(x)-f(x)|^p dx \leqq C_p\left\{\mathrm{E}\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|^p dx+\int|\mathrm{E}\,\hat{f}(x)-f(x)|^p dx\right\},$$

so that

$$\text{(3.2)}\qquad \mathrm{E}\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|^p dx \leqq \mathrm{E}^{1/2}[\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|]^{2(p-1)}\,\mathrm{E}^{1/2}\left[\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|^2 dx\right],$$

where it follows from Theorem 1 that $\mathrm{E}^{1/2}[\sup_x|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|]^{2(p-1)}\leqq C(na_n^m)^{-(p-1)/2}$, and also $\mathrm{E}\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|^2dx=\int\mathrm{Var}\,(\hat{f}(x))dx=(na_n^m)^{-1}\cdot\int k^2(u)du+O((na_n^m)^{-1})$. Hence

$$\text{(3.3)}\qquad ||\hat{f}-\mathrm{E}\,\hat{f}||_p=\left\{\mathrm{E}\int|\hat{f}(x)-\mathrm{E}\,\hat{f}(x)|^p dx\right\}^{1/p}=O((na_n^m)^{-1/2}).$$

Let us evaluate $||\mathrm{E}\,\hat{f}-f||_p^p=\int|\mathrm{E}\,\hat{f}(x)-f(x)|^p dx$. Note that

$$\text{(3.4)}\qquad \begin{aligned}\mathrm{E}\,\hat{f}(x)-f(x)&=\int[f(x-a_nu)-f(x)]k(u)du\\ &=\int\Big[\sum_{j=1}^{M-1}\frac{1}{j!}f^{(j)}(x;\,-a_nu)\\ &\qquad+\frac{1}{M!}\int_0^1 f^{(M)}(x-\theta a_nu;\,-a_nu)d\theta\Big]k(u)du,\end{aligned}$$

where $f^{(1)}(x;t)=\sum_{i=1}^m\frac{\partial f(x)}{\partial x_i}t_i$, $f^{(2)}(x;t)=\sum_{i=1}^m\sum_{j=1}^m\frac{\partial^2 f(x)}{\partial x_i\partial x_j}t_it_j$, $\cdots$ generally $f^{(j)}(x;t)=\sum_{i_1=1}^m\cdots\sum_{i_j=1}^m\frac{\partial^j f(x)}{\partial x_{i_1}\cdots\partial x_{i_j}}t_{i_1}\cdots t_{i_j}$. Note that if $\int u_{i_1}\cdots u_{i_j}k(u)du=0$ for all $i_1,\cdots,i_j=1,\cdots,m$ and $j=1,2,\cdots,M-1$, then we get

$$\text{(3.5)}\qquad \mathrm{E}\,\hat{f}(x)-f(x)=\frac{1}{M!}\iint_0^1 k(u)f^{(M)}(x-\theta a_nu;\,-a_nu)d\theta du.$$

But it $\int|u_{i_1}|\cdots|u_{i_M}|k(u)du<\infty$ for all $i_1,i_2,\cdots,i_M=1,2,\cdots,m$ and if

$f(x)$ has all of its partial derivatives of order M in L_p, $1 \leqq p < \infty$, then

$$(3.6)\quad \int |\mathrm{E}\,\hat{f}(x) - f(x)|^p dx \leqq \int (M!)^{-p} \left| \int k(u) \int_0^1 f^{(M)}(x - \theta a_n u;\, -a_n u) d\theta du \right|^p dx .$$

We shall show that $a_n^{-mp} \int \left| \int k(u) \int_0^1 f^{(M)}(x - \theta a_n u;\, -a_n u) d\theta du \right|^p dx$ has a finite limit. Let $D(i_1, \cdots, i_M; f)(x) = \partial^M f(x) / \partial x_{i_1} \cdots \partial x_{i_M}$. Thus

$$(3.7)\quad f^{(M)}(x - \theta a_n u;\, -a_n u) = \sum_{i_1=1}^{m} \cdots \sum_{i_M=1}^{m} D(i_1, \cdots, i_M; f)(x - \theta a_n u) u_{i_1} \cdots u_{i_M} \cdot a_n^M (-1)^M .$$

Hence

$$(3.8)\quad a_n^{-mp} \int \left| k(u) \int_0^1 f^{(M)}(x - \theta a_n u;\, -a_n u) d\theta du \right|^p dx \leqq C_p \sum_{i_1=1}^{m} \cdots \sum_{i_M=1}^{m} \int \left| k(u) \left(\int_0^1 D(i_1, \cdots, i_M; f)(x - \theta a_n u) d\theta \right) u_{i_1} \cdots u_{i_M} \right|^p \times du dx ,$$

which converges to $C_p \sum_{i_1=1}^{m} \cdots \sum_{i_M=1}^{m} \left(\int |u_{i_1} \cdots u_{i_M}|^p k(u) du \right) \int |D^p(i_1, \cdots, i_M; f)(x)| \cdot dx < \infty$. We therefore proved that $\mathrm{E}\,\|\hat{f} - f\|_p = O((n a_n^m)^{-1/2} + a_n^{mM})$, $1 \leqq p < \infty$, thus for $a_n = n^{-1/(2M+1)m}$ we get that $\mathrm{E}\,\|\hat{f} - f\|_p = O(n^{mM/(2M+1)m})$. When $p = \infty$ we proceed exactly as above but assuming that all partial derivatives of $f(x)$ of order M are bounded and we also get, for $a_n = n^{-1/(M+1)m}$, that $\mathrm{E}\,\|\hat{f} - f\|_\infty = O(n^{-M/(2M+1)})$; collecting the above we have proved the following theorem.

THEOREM 3. *Suppose that $k(u)$ satisfies, $\int u_{i_1} \cdots u_{i_j} k(u) du = 0$ for all $i_1, \cdots, i_j = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, M-1$, and $\int |u_{i_1}| \cdots |u_{i_M}| k(u) du < \infty$ for all $i_1, \cdots, i_M = 1, 2, \cdots, m$. Assume that all partial derivatives of order M or less exist and take $a_n = n^{-1/m(2M+1)}$.*

(i) *If the partial derivatives of order M of f are all in L_p, for $1 \leqq p < \infty$, then,*

$$(3.9)\quad \mathrm{E}\,\|\hat{f} - f\|_p = O(n^{-M/(2M+1)}) .$$

(ii) *If the partial derivatives of order M of f are bounded, then*

$$(3.10)\quad \mathrm{E}\,\|\hat{f} - f\|_\infty = O(n^{-M/(2M+1)}) .$$

In the next theorem we discuss the rates of convergence in the strong L_p-consistency. Our main tool is the following result of Kuelbs [4]. Let B denote a real vector space, $\mathscr{B}$ a σ-field of subsets of B, and $\|\cdot\|$ a seminorm on B. The triplet $(B, \mathscr{B}, \|\cdot\|)$ is said to be a *linear measurable space* if the following conditions are satisfied:

(i) Addition and scalar multiplication are $\mathcal{B}$-measurable operations.
(ii) For all $t \geqq 0$, the set $\{x \in B: ||x|| \leqq t\}$ is $\mathcal{B}$-measurable.
(iii) There exists a subset F of the $\mathcal{B}$-measurable linear functionals on B such that

$$||x|| = \sup_{f \in F} |f(x)|, \qquad x \in B. \tag{3.11}$$

Examples of linear measurable spaces include the R^m Euclidean space. Kuelbs [4] proved the following:

THEOREM 4. *Let $(B, \mathcal{B}, ||\cdot||)$ be a linear measurable space and assume that $X_1, X_2, \cdots$ are independent $(B, \mathcal{B})$-valued random variables such that* $\mathrm{E}\, f(X_j)=0$ *for all $f \in F$ and all $j \geqq 1$,* $\sup_{j \geqq 1} \mathrm{E}\,[\exp(\beta ||X_j||^2)] < \infty$ *for all $\beta > 0$, and for some sequence of positive constants $\{b_j\}$ we have; (i) $\sigma_n^2 = \sum_{i=1}^n b_i^2 \to \infty$ as $n \to \infty$, (ii) $b_n^2/\sigma_n^2 \to 0$ as $n \to \infty$, and* (iii) $\left\{\sum_{j=1}^n b_j X_j / \sigma_n\right\}_{n=1}^\infty$ *is bounded inprobability. If $S_n = \sum_{j=1}^n b_j X_j$ and $C_n = (2\sigma_n^2 \ln \ln \sigma_n^2)^{1/2}$, then for all $\beta > 0$,*

$$\mathrm{E}\,[\exp\{\beta \sup_n ||S_n/C_n||^2\}] < \infty. \tag{3.12}$$

Note that if $\Delta_n = \sup_x |F_n(x) - F(x)|$ denote the empirical process and we set $M_n = \sup_n (n/2 \ln \ln n)^{1/2} \Delta_n$, then Theorem 4 applies and we have for all $\beta > 0$, $\mathrm{E}\,[\exp(\beta M_n^2)] < \infty$. This fact is utilized in developing the following result.

THEOREM 5. *Suppose that $k(u)$ satisfies the conditions imposed on it in Theorem 2. Assume further that f has all partial derivatives of order M or less and take $a_n = (n/\ln \ln n)^{-1/2m(M+1)}$.*
(i) *If the partial derivatives of order M of f are all in L_p, $1 \leqq p < \infty$, then*

$$||\hat{f} - f||_p = O((n/\ln \ln n)^{-M/2(M+1)}), \tag{3.13}$$

and

$$\lim_{n \to \infty} \mathrm{E}\,[\exp\{\beta ||\hat{f} - f||_p\} = 1. \tag{3.14}$$

(ii) *If the partial derivatives of order M of f are all bounded, then* (3.13) *and* (3.14) *hold for $p = \infty$.*

PROOF. It follows from Theorem 2 that for any $1 \leqq p \leqq \infty$

$$||\hat{f} - \mathrm{E}\,\hat{f}||_p \leqq C a_n^{-m} \Delta_n \leqq C(\ln \ln n / n a_n^{2m})^{1/2}. \tag{3.15}$$

But also it follows from Theorem 3 that for any $1 \leqq p \leqq \infty$

$$(3.16) \qquad ||\mathrm{E}\,\hat{f}-f||_p=O(a_n^{mM})\,.$$

From (3.15) and (3.16), (3.13) follows by taking $a_n=(n/\ln\ln n)^{-1/2m(M+1)}$. Next, (3.14) also follows from (3.15) and (3.16) in light of the remark immediately following Theorem 4.

We conclude this section by noting that the rates obtained in Theorems 3 and 5 are independent from the dimensionality of $f(x)$, however, the best choice of a_n leading to these estimates does.

Acknowledgement

The author is indebted to the referee for his illuminating remarks which resulted in an improved paper. This research is supported in part by a grant from the University of Petroleum and Minerals.

UNIVERSITY OF PETROLEUM AND MINERALS

REFERENCES

[1] Cacoullos, T. (1966). Estimation of a multivariate density, *Ann. Inst. Statist. Math.*, 18, 179-189.

[2] Dharmadhikari, S. W., Fabian, V. and Jogdeo, K. (1968). Bounds on the moments of martingales, *Ann. Math. Statist.*, **39**, 1719-1723.

[3] Fryer, M. J. (1977). A review of some nonparametric methods of density estimation, *J. Inst. Math. Appl.*, **20**, 335-354.

[4] Kuelbs, J. (1978). Some exponential moments of sums of independent random variables, *Trans. Amer. Math. Soc.*, **240**, 145-162.

[5] Leadbetter, M. R. (1963). On the nonparametric estimation of probability densities, *Tech. Report* No. 11, Research Triangle Institute.

[6] Loeve, M. (1963). *Probability Theory*, Von Nostrand.

[7] Matusita, K. (1955). Decision rules based on the distance for the problem of fit, two samples, and estimation, *Ann. Math. Statist.*, **26**, 631-640.

[8] Nadaraya, E. A. (1965). On nonparametric estimates of density functions and regression curves, *Theory Prob. Appl.*, **10**, 186-190.

[9] Parzen, E. (1962). On estimation of a probability density and mode, *Ann. Math. Statist.*, **33**, 1065-1076.

[10] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function, *Ann. Math. Statist.*, **27**, 832-837.

[11] Rüschendorf, L. (1977). Consistency of estimators for multivariate density functions and for the mode, *Sankhyā*, A, **39**, 243-350.

[12] Schuster, E. F. (1970). Note on the uniform convergence of density estimates, *Ann. Math. Statist.*, **41**, 1347-1348.

[13] Van Ryzin, J. (1969). On strong consistency of density estimates, *Ann. Math. Statist.*, **41**, 1765-1772.