Z. Wahrscheinlichkeitstheorie verw. Gebiete
61, 405-415 (1982)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Weak and Strong Uniform Consistency of Kernel Regression Estimates*

Y.P. Mack[1] and B.W. Silverman[2]

[1] University of California, Division of Statistics, Davis, CA 95616, U.S.A.
[2] University of Bath, School of Mathematics, Claverton Downs, Bath, BA2 7AY, United Kingdom

Summary. We study the estimation of a regression function by the kernel method. Under mild conditions on the "window", the "bandwidth" and the underlying distribution of the bivariate observations $\{(X_i, Y_i)\}$, we obtain the weak and strong uniform convergence rates on a bounded interval. These results parallel those of Silverman (1978) on density estimation and extend those of Schuster and Yakowitz (1979) and Collomb (1979) on regression estimation.

I. Introduction

Let (X, Y), (X_i, Y_i), $i=1,2,\ldots$ be i.i.d. bivariate random variables with common joint distribution $F(x,y)$ and joint density $f(x,y)$. Let $g(x)$ be the marginal density of X and let $r(x)=E(Y|X=x)$ be the regression of Y on X. Nadaraya (1964) and Watson (1964) independently proposed nonparametric estimators of $r(x)$ based on the kernel method as introduced by Rosenblatt (1956) for density estimation. Specifically, they have the form

$$r_n(x)=h_n(x)/g_n(x), \tag{1}$$

where

$$h_n(x)=\frac{1}{n b_n}\sum_{j=1}^{n}\delta\left(\frac{x-X_j}{b_n}\right)\cdot Y_j$$

and

$$g_n(x)=\frac{1}{n b_n}\sum_{j=1}^{n}\delta\left(\frac{x-X_j}{b_n}\right).$$

Here δ is a kernel function and $\{b_n\}$ is a sequence of "bandwidths" tending to zero as n tends to infinity. Watson (1964) gave some heuristic analysis of (1) in

* This research was carried out in part while the authors were guests at the University of Heidelberg, Germany, under the sponsorship of the Sonderforschungsbereich 123 in the summer of 1980

0044-3719/82/0061/0405/$02.20

conjunction with another class of nonparametric regression estimates which are the forerunners of the nearest neighbor method (see Royall (1966), Stone (1977) and Mack (1981), for instance). Rosenblatt (1969) obtained the bias, variance and asymptotic distribution of (1). Analogous results for the multivariate case were obtained by Collomb (1977). Schuster (1972) demonstrated the multivariate normality at a finite number of distinct points. Major (1973) studied a variant of r_n (where g is estimated by a histogram) and obtained the asymptotic distribution of maximum deviations as well as rate of strong uniform consistency on a finite interval. Nadaraya (1974) considered the limiting distribution of the weighted quadratic functional given by $\int [r_n(x) - r(x)]^2 \cdot g_n(x)^2\,dx$. Noda (1976) proved the pointwise strong consistency of (1) and also derived the rate of convergence of the *MSE* at a continuity point of $r(x)$. Konakov (1977) investigated the behavior of the weighted quadratic deviation $\int [r_n(x) - r(x)]^2 \cdot p(x) \cdot g_n(x)^2\,dx$ for a class of weight functions p.

Recently, Schuster and Yakowitz (1979) derived uniform convergence bounds and uniform consistency on a finite interval for r_n. Johnston (1979) and Wandl (1980) studied the global deviation along the same lines as Bickel and Rosenblatt (1973), and Revesz (1979) considered similar questions for variants of kernel and nearest neighbour regression estimates. Under quite different conditions on the joint distribution of (X, Y) from those of the present paper, and without considering rates of consistency, Collomb (1979) gave necessary and sufficient conditions on the bandwidth for strong uniform consistency of r_n; his main condition is $n^{-1} b_n^{-1} \log n \to 0$. Independently of us, Wandl (1980) has obtained rates of uniform consistency. Though these rates contain an exact scale constant, the conditions on the bandwidth are much more restrictive than ours and in addition the marginal distribution of Y is assumed to have bounded support; thus, although Wandl's results are deep, even the standard case of regression with normal errors is not covered. Major (1973) obtained similar results for a different class of estimators.

In the present paper, we derive the weak and strong consistency of r_n on a bounded interval, together with rates of convergence. Our conditions involve only a moment condition on Y and mild conditions on the bandwidth and the smoothness of the kernel. Under these assumptions, we show that the weak and strong uniform convergence rates are $O(n^{-\frac{1}{2}} b_n^{-\frac{1}{2}} \log(1/b_n))$; this is the rate shown by Wandl (1980) to be best possible under much more restrictive assumptions. The results of Schuster and Yakowitz (1979) yield a rate of uniform convergence in probability slightly slower than $O(n^{-\frac{1}{2}} b_n^{-1})$; their bounds are not good enough to allow *any* rate of strong convergence to be obtained by the use of the Borel-Cantelli lemma.

Note that although only the two-dimensional case is considered here, the technique can be extended to the higher dimensional case, with different rates of convergence. These rates, however, will not be compatible with the univariate case since it is still unknown whether the strong uniform approximation of the multivariate empirical process by a multivariate Brownian bridge (a device which is necessary in our analysis) has a compatible rate as in the univariate or bivariate case. Nevertheless, one can proceed as in Rosenblatt (1976) by appealing to the results of Csörgő and Révész (1975) and Révész

(1976) to derive some multivariate regression estimation results for the multivariate case.

As footnotes to the above discussion, first, the idea of invariance principle in nonparametric regression has been exploited earlier by Bhattacharya (1976), although the form of the estimate considered there was based on induced order statistics (or concomitants); second, we would like to mention some related works on density estimation where strong uniform consistency was proved without the invariance principle. These include, among others, Bosq (1970), Deheuvels (1974), Bertrand-Retali (1974), Geffroy (1974), Reiss (1975), Bleuez and Bosq (1976), Hall (1981) and Stute (1982).

The main results of this paper are given in Sect. 3 at the end; in the next section we give the technical preliminaries to these results.

II. Some Preliminaries

Let

$$h(x)=\int y f(x,y)\,dy$$

and write (suppressing the arguments of the functions)

$$r_n - r = A + B + C, \tag{2}$$

where

$$A = g_n^{-1}\cdot(h_n - E\,h_n),$$

$$B = Eh_n\cdot(g_n\cdot E\,g_n)^{-1}\cdot(E\,g_n - g_n),$$

$$C = (E\,g_n)^{-1}\cdot(E\,h_n - h) + h\cdot(g\cdot E\,g_n)^{-1}\cdot(g - E\,g_n).$$

In order to study the uniform convergence of $r_n(x)$ over a bounded interval J on which g is bounded away from zero, by using (2) together with the work of Silverman (1978) on the weak and strong consistency of $g_n(x)$, it is enough to study that of $h_n(x)$. In fact, most of our arguments follow the same fashion of that paper except for a truncation applied to $h_n(x)$ at the beginning so that we can carry out some integration by parts later. The outline of the present paper is thus organized as follows:

(i) truncation;
(ii) strong approximation;
(iii) uniform convergence with rates.

Throughout this discussion, unless otherwise stated, sup, inf and $\int$ will be taken over the entire real line. We assume $g(x)$ and $h(x)$ are continuous on $\mathbb{R}^1$. Other conditions on g and h will be imposed as the occasion demands. The kernel function δ is assumed to satisfy condition

(C1) (a) δ is uniformly continuous with modulus of continuity w_δ and of bounded variation $V(\delta)$;
(b) δ is absolutely integrable w.r.t. Lebesgue measure on the line;
(c) $\delta(x)\to 0$ as $|x|\to\infty$;
(d) $\int |x\log|x||^{\frac{1}{2}}\,|d\delta(x)|<\infty$.

Let B_n define an increasing sequence of positive numbers such that $B_n \to \infty$ as $n \to \infty$. We introduce also the truncated estimate

$$h_n^B(x) = \frac{1}{n\,b_n} \sum_{j=1}^{n} \delta\left(\frac{x-X_j}{b_n}\right) \cdot Y_j \cdot I\{|Y_j| < B_n\}$$
$$= \frac{1}{b_n} \iint_{|y|<B_n} \delta\left(\frac{x-t}{b_n}\right) y\, dF_n(t, y)$$

where $F_n(\cdot, \cdot)$ is the empirical distribution of the (X_j, Y_j)'s, and I is the indicator function on a set.

Our first result deals with the truncation error in replacing h_n by h_n^B.

Proposition 1. *Suppose, for some $s>0$, $E|Y|^s < \infty$*

$$\sup_x \int |y|^s f(x, y)\, dy = M_s < \infty.$$

Provided δ is absolutely integrable and B_n is an increasing sequence for which $\Sigma_n B_n^{-s}$ converges,

$$\sup_x |h_n(x) - h_n^B(x) - E\{h_n(x) - h_n^B(x)\}| = O(B_n^{1-s})$$

with probability 1 (w.p.1).

Proof. Since $P(|Y_n| > B_n) \leqq B_n^{-s} E|Y|^s$, it follows that, w.p.1, $|Y_n| \leqq B_n$ for all sufficiently large n, and hence, since B_n is increasing, by elementary real analysis, that for all sufficiently large n,

$$|Y_j| \leqq B_n \qquad \text{for all } j \leqq n.$$

This implies that $\sup |h_n(x) - h_n^B(x)|$ is eventually zero w.p.1.

It remains to bound the expectation term. Using the fact that, by a standard argument,

$$\sup_x \int_{|y| \geqq B_n} |y|\, f(x, y)\, dy \leqq M_s B_n^{1-s},$$

it follows that

$$E|h_n(x) - h_n^B(x)| \leqq \frac{1}{b_n} \iint_{|y| \geqq B_n} \left|\delta\left(\frac{x-t}{b_n}\right)\right| |y|\, f(t, y)\, dy\, dt$$
$$\leqq \int |\delta(\xi)|\, d\xi \cdot \sup_\tau \int_{|y| \geqq B_n} |y|\, f(\tau, y)\, dy$$
$$= O(B_n^{1-s}) \quad \text{w.p.1.}$$

Combining the results of the last two paragraphs completes the proof.

Next, we give a decomposition of the truncated estimate $h_n^B(x)$.

Proposition 2. *On a rich enough probability space, there exists a version B^0 of the two-dimensional Brownian bridge such that*

$$h_n^B(x) = E[h_n^B(x)] + n^{-\frac{1}{2}} \rho_n(x) + \varepsilon_n(x),$$

where

$$\rho_n(x) = b_n^{-1} \iint_{|y|<B_n} \delta\left(\frac{x-t}{b_n}\right) y \, dB^0(T(t,y))$$

with T: $\mathbb{R}^2 \to [0,1]^2$ *denoting the transformation* (see Rosenblatt (1952)) *defined by*

$$T(x,y) = (F_X(x), F_{Y|X}(y|x)),$$

F_X *being the marginal d.f. of* X *and* $F_{Y|X}$ *the conditional d.f. of* Y *given* X, *and where*

$$\varepsilon_n(x) = -n^{-\frac{1}{2}} b_n^{-1} \int \Delta_n(t) \, d_t \, \delta\left(\frac{x-t}{b_n}\right) \tag{3}$$

with

$$\Delta_n(t) = \int_{|y|<B_n} y \, d_y [Z_n(t,y) - B^0(T(t,y))],$$

$Z_n(t,y) = n^{\frac{1}{2}}[F_n(t,y) - F(t,y)]$ *is the two-dimensional empirical process, and*

$$\sup |\varepsilon_n(t)| = O[B_n (n b_n)^{-1} (\log n)^2] \quad \text{w.p.1.} \tag{4}$$

Proof. Write

$$h_n^B(x) - E\, h_n^B(x) = n^{-\frac{1}{2}} b_n^{-1} \iint_{|y|<B_n} \delta\left(\frac{x-t}{b_n}\right) y \, dZ_n(t,y)$$

$$= n^{-\frac{1}{2}} b_n^{-1} \int \delta\left(\frac{x-t}{b_n}\right) dU_n^B(t)$$

where

$$U_n^B(t) = \int_{|y|<B_n} y \, d_y Z_n(t,y).$$

Employing the strong approximation result of Tusnády (1977) for a two-dimensional empirical process, we can find a suitable probability space supporting both Z_n and a two-dimensional Brownian bridge B^0 such that

$$\sup_{t,y} |Z_n(t,y) - B^0(T(t,y))| = O(n^{-\frac{1}{2}} (\log n)^2) \quad \text{w.p.1.}$$

Hence (3) follows by using an integration by parts. Replacing Z_n by B^0 after a change of variables, and keeping track of the error in the approximation, we obtain (4).

Next, we state the strong and weak uniform convergence rates of the Gaussian process ρ_n. For convenience, we let

$$\alpha_n = b_n^{\frac{1}{2}} \left(\log \frac{1}{b_n}\right)^{-\frac{1}{2}}.$$

Proposition 3. *Suppose* $M = \sup_x \int y^2 f(x,y)\, dy$ *is finite,* δ *satisfies condition* (C1). *Then*

$$\alpha_n \sup |\rho_n(x)| = O_p(1).$$

If, in addition, $\Sigma_n b_n^\lambda < \infty$ *for some* $\lambda > 0$, *then*

$$\alpha_n \sup |\rho_n(x)| = O(1) \quad \text{w.p.1.} \tag{5}$$

Proof. Since the arguments used here are analogous with those used in Silverman (1976, 1978), we give only the essentials. Define

$$U_n^*(t) = \int_{|y|<B_n} y \, dB^0(T(t, y)).$$

Note that U_n^* has the same covariance structure as U_n^B, which is (for $s<t$)

$$\operatorname{cov}(Y \cdot I\{X<t, \ |Y|<B_n\}, \ Y \cdot I\{X<s, \ |Y|<B_n\}) = v_n^B(s) - \mu_n^B(s) \cdot \mu_n^B(t),$$

where

$$v_n^B(\cdot) = \int_{|y|<B_n} y^2 \, d_y F(\cdot, y)$$

and

$$\mu_n^B(\cdot) = \int_{|y|<B_n} y \, d_y F(\cdot, y).$$

Hence

$$\begin{aligned} E[U_n^*(s) - U_n^*(t)]^2 &= v_n^B(t) - v_n^B(s) - [\mu_n^B(s) - \mu_n^B(t)]^2 \\ &\leqq v_n^B(t) - v_n^B(s) \\ &= \int_{|y|<B_n} \int_s^t y^2 f(x, y) \, dx \, dy \\ &\leqq \int_{-\infty}^{\infty} \int_s^t y^2 f(x, y) \, dx \, dy \\ &= v(t) - v(s), \end{aligned} \tag{6}$$

where

$$v(\cdot) = \int y^2 \, d_y F(\cdot, y).$$

Thus under the assumption that $\sup_x \int y^2 f(x, y) \, dy = M < \infty$, we have that for all x,

$$v'(x) = \int y^2 f(x, y) \, dy \leqq M < \infty.$$

Therefore, for $0 < \tau < \sigma^2 = E(Y^2) = v(\infty)$, $v^{-1}(\tau)$ can be defined so that $v^{-1}\{v(\tau)\} = \tau$. Define

$$V_n(\tau) = U_n^*(v^{-1}(\tau)).$$

It follows from (6) that $U_n^*(s) = V_n\{v(s)\}$ for all s and that

$$E[V_n(\tau) - V_n(\tau')]^2 \leqq |\tau - \tau'|.$$

Let θ be the modulus of continuity of V_n. Then the trivial generalization of Garsia (1970), given by Silverman (1976), Lemma 2, implies that there exists a random variable A with $EA \leqq 4\sqrt{2}\sigma^4$ such that $\theta(\varepsilon)$ is majorized by

$$16(\log A)^{\frac{1}{2}} \varepsilon^{\frac{1}{2}} + 16\sqrt{2} q(\varepsilon),$$

where

$$q(\varepsilon) = \int_0^\varepsilon \frac{1}{2} \left[\frac{1}{r} \log \frac{1}{r}\right]^{\frac{1}{2}} dr.$$

Hence,

$$\begin{aligned}|\rho_n(x)| &= \left|\frac{1}{b_n}\int \delta\left(\frac{x-t}{b_n}\right) dU_n^*(t)\right| \\ &= \left|\frac{1}{b_n}\int [U_n^*(x-b_n\tau)-U_n^*(x)]\, d\delta(\tau)\right| \\ &\leqq \frac{1}{b_n}\int |V_n(v(x-b_n\tau))-V_n(v(x))|\cdot|d\delta(\tau)| \\ &\leqq \frac{1}{b_n}\int \theta(M b_n|\tau|)\cdot|d\delta(\tau)|.\end{aligned}$$

Thus

$$\begin{aligned}\alpha_n \sup|\rho_n(x)| \leqq\; & 16(\log A)^{\frac{1}{2}} M^{\frac{1}{2}}\left(\log\frac{1}{b_n}\right)^{-\frac{1}{2}}\cdot\int|\tau|^{\frac{1}{2}}|d\delta(\tau)| \\ & +16\sqrt{2}\, b_n^{-\frac{1}{2}}\left(\log\frac{1}{b_n}\right)^{-\frac{1}{2}}\int q(M b_n|\tau|)\,|d\delta(\tau)|. \qquad (7)\end{aligned}$$

Using arguments similar to Silverman (1978), pp. 180–181, the second term of (7) tends to

$$16\sqrt{2}M^{\frac{1}{2}}\int|\tau|^{\frac{1}{2}}|d\delta(\tau)|$$

by conditions (a) and (d) of (C1); and the first term of (7) is $O_p(1)$ if $b_n\to 0$; and is $O(1)$ w.p.1 if $\Sigma_n b_n^\lambda<\infty$ for some $\lambda>0$.

The final result of this section combines Propositions 1, 2 and 3 to give bounds on the behaviour of $h_n - E h_n$.

Proposition 4. *Suppose the conditions on δ and f of Propositions* 1, 2 *and* 3 *hold, and that as $n\to\infty$, $b_n\to 0$ and $n^\eta b_n\to\infty$ for some $\eta<1-s^{-1}$. Then*

$$\sup|h_n(x)-E\,h_n(x)| = o_p(1). \qquad (8)$$

If, in addition, $n^{2\eta-1}b_n\to\infty$ then

$$\sup|h_n(x)-E\,h_n(x)| = O_p(n^{-\frac{1}{2}}\alpha_n^{-1}). \qquad (9)$$

If the condition $b_n\to 0$ is replaced by $\Sigma_n b_n^\lambda<\infty$ for some $\lambda>0$, then the probability orders of magnitude in (8) *and* (9) *can be replaced by orders of magnitude* w.p.1.

Proof. Using the notation of Proposition 1, 2 and 3, let

$$E_1=\sup|\varepsilon_n(x)|,$$

$$E_2=\sup|n^{-\frac{1}{2}}\rho_n(x)|,$$

and

$$E_3=\sup|h_n(x)-h_n^B(x)-E\{h_n(x)-h_n^B(x)\}|.$$

Then $\sup|h_n-E h_n|$ is dominated by $E_1+E_2+E_3$, and bounds for the E_i are given in the propositions above. Suppose first that $b_n\to 0$ and $n^\eta b_n\to\infty$. Set $\delta=\frac{1}{3}(1-s^{-1}-\eta)$ and $B_n=n^{\frac{1}{s}+\delta}$. Then, applying Propositions 1 and 2, with probability 1

$$E_1 = o(n^{-(s-1)/s} b_n^{-1} n^{\delta} (\log n)^2)$$
$$= o(n^{-\eta} b_n^{-1}) \tag{10}$$

and

$$E_3 = o(n^{-(s-1)/s}) = o(n^{-\eta}). \tag{11}$$

Since, from Proposition 3, $E_2 = O_p(n^{-\frac{1}{2}} \alpha_n^{-1})$ it follows that $E_2 = o_p(1)$ and hence, using (10) and (11), that (8) holds.

From (10) and (11) it follows that, w.p.1.,

$$\{n^{\frac{1}{2}} \alpha_n (E_1 + E_3)\}^2 = o(n^{1-2\eta} b_n^{-1})$$

and hence, using results for E_2, that (9) holds under the specified condition. The final remark of the proposition follows immediately by appealing to (5) in Proposition 3.

III. Main Results

In what follows, set

$$\theta_n = \left(\frac{1}{n b_n} \log \frac{1}{b_n}\right)^{\frac{1}{2}}$$

and assume the following condition on f, g and h:

(C2) (a) $E|Y|^s < \infty$ and $\sup_x \int |y|^s f(x, y)\, dy < \infty$, $s \geqq 2$;

(b) f, g and h are continuous on an open interval containing J.

Note that $E|Y|^s < \infty$ is equivalent to the integrability w.r.t. x of $\int |y|^s f(x, y)\, dy$, and so the boundedness of this quantity is a very mild additional condition. It will be seen that under appropriate conditions on b_n, the weak and strong uniform convergence rates of the regression estimate over a suitable bounded interval J are $O(\theta_n)$. The first two lemmas are known results in the literature.

Lemma 1 (Silverman, 1978). *Suppose δ satisfies condition* (C1) *and g is uniformly continuous. Provided $b_n \to 0$ and $n^{1-\varepsilon} b_n \to \infty$ for some $\varepsilon > 0$,*

$$\sup_x |g_n(x) - g(x)| = o(1) \quad \text{w.p.1}$$

and

$$\theta_n^{-1} \sup_x |g_n(x) - E g_n(x)| = O_p(1).$$

If also $\Sigma_n b_n^{\lambda} < \infty$ for some $\lambda > 0$ then

$$\theta_n^{-1} \sup_x |g_n(x) - E g_n(x)| = O(1) \quad \text{w.p.1.}$$

Lemma 2 (Parzen, 1962). *Suppose δ satisfies conditions* (C1) *and ϕ is continuous and absolutely integrable. Let ϕ_n be the convolution of ϕ and $\delta_n(u) = \delta(u/b_n)/b_n$. Then*

$$\sup_J |\phi_n(x) - \phi(x)| = o(1).$$

We are now ready to state our main results.

Theorem A. *Suppose δ satisfies conditions* (C1) *and f satisfies conditions* (C2). *Suppose J is a bounded interval on which g is bounded away from zero. Suppose that $\Sigma_n b_n^\lambda < \infty$ for some $\lambda > 0$ and that $n^\eta b_n \to \infty$ for some $\eta < 1 - s^{-1}$. Then*

$$\sup_J |r_n(x) - r(x)| = o(1) \quad \text{w.p.1.}$$

Theorem B. *Suppose the conditions of Theorem A for δ and J hold, and that f satisfies conditions* (C2) *for some $s > 2$. Suppose $n^{2\eta - 1} b_n \to \infty$ for some $\eta < 1 - s^{-1}$. Furthermore, suppose g and h have bounded second derivatives, $b_n^2 = o(\theta_n)$, and $\delta(u) = \delta(-u)$, then*

$$\theta_n^{-1} \sup_J |r_n(x) - r(x)| = O_p(1); \tag{12}$$

and if $\Sigma_n b_n^\lambda < \infty$ for some $\lambda > 0$, then

$$\theta_n^{-1} \sup_J |r_n(x) - r(x)| = O(1) \quad \text{w.p.1.} \tag{13}$$

The proofs of the theorems involve in the first step a decomposition of r_n according to (2). For Theorem A, we apply Proposition 4 and the lemmas above by noting that h is absolutely integrable by (C2). For Theorem B, we apply Proposition 4 and deal with the bias in the usual way by Taylor expansion.

Note that if we put $b_n = n^{-\alpha}$ for some fixed $\alpha > 0$, then the condition for strong uniform consistency is $0 < \alpha < 1 - s^{-1}$, while for weak and strong uniform consistency with rate θ_n, the condition is $\frac{1}{5} < \alpha < 1 - \frac{2}{s}$. The last condition requires that Y has s-th absolute moment for $s > \frac{5}{2}$. However, if Y is bounded a.s., then the condition becomes $\frac{1}{5} < \alpha < 1$, which is still considerably weaker than the conditions imposed on the bandwidth by Major (1973) and Wandl (1980).

References

Bertrand-Retali, M.: Convergence uniforme stochastique d'un estimator d'une densité de probabilité dans $\mathbb{R}^s$. C.R. Acad. Sci. Paris **278**, 1449–1452 (1974)

Bhattacharya, P.K.: An invariance principle in regression analysis. Ann. Statist. **4**, 621–624 (1976)

Bickel, P.J., Rosenblatt, M.: On some global measures of the deviations of density function estimates. Ann. Statist. **1**, 1071–1095 (1973)

Bleuez, J., Bosq, D.: Conditions nécessaires et suffisantes de convergence pour une class d'estimateurs de la densité. C.R. Acad. Sci. Paris **282**, 636–666 (1976)

Bosq, D.: Contribution à la théorie de l'estimation fonctionnelle. Publications de l'Institut de Statistique de l'Université de Paris, Vol. 19, fasc. 2 et 3 (1970)

Cheng, K.F., Taylor, R.L.: On the uniform complete convergence of estimates for multivariate density functions and regression curves. Ann. Inst. Statist. Math. **32**, 187–199 (1980)

Collomb, G.: Estimation non-paramétrique de la régression par la methode du noyau. Thèse, Université Paul Sabatier, Toulouse (1976)

Collomb, G.: Quelques propriétés de la méthode du noyau pour l'estimation non-paramétrique de la régression en un point fixé. C.R. Acad. Sci. Paris **285**, 289–292 (1977)

Collomb, G.: Conditions nécessaires et suffisantes de convergence uniforme d'un estimateur de la régression, estimation des dérivées de la regression, C.R. Acad. Sci. Paris **288**, 161–164 (1979)
Collomb, G.: Estimation non paramétrique de la régression: revue bibliographique. ISR **49**, 75–93 (1981)
Csörgő, M., Révész, P.: A new method to prove Strassen type laws of invariance principle II. Z. Wahrscheinlichkeitstheorie verw. Gebiete **31**, 261–269 (1975)
Deheuvels, P.: Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre et uniforme presque sûre des estimateurs de la densité. C.R. Acad. Sci. Paris **278**, 1217–1220 (1974)
Devroye, L.P.: The uniform convergence of the Nadaraya-Watson regression function estimate. Canad. J. Statist. **6**, 179–191 (1979)
Garsia, A.M.: Continuity properties of Gaussian processes with multidimensional time parameter. Proc. Sixth Berkeley Sympos. Math. Statist. Probab. **2**, 369–374. Univ. of California Press (1970)
Geffroy, J.: Étude de la convergence du régressogramme. Séminarie de Statistique Mathématique de l'Institut de Mathématique de l'Université de Paris, année 1974–1975 (1975)
Hall, P.: Laws of the iterated logarithm for nonparametric density estimators. Z. Wahrscheinlichkeitstheorie verw. Gebiete **50**, 47–61 (1981)
Johnston, G.J.: Smooth nonparametric regression analysis. Ph.D. dissertation, U. of No. Carolina at Chapel Hill (1979)
Komlós, J., Major, P., Tusnády, G.: An approximation of partial sums of independent random variables, and the sample distribution function I. Z. Wahrscheinlichkeitstheorie verw. Gebiete **32**, 111–131 (1975)
Konakov, V.D.: On a global measure of deviation for an estimate of the regression line. Theor. Probab. Appl. **22**, No. 4, 858–868 (1977)
Mack, Y.P.: Local properties of $k-NN$ regression estimates. SIAM. J. Alg. Disc. Meth. **2**, 311–323 (1981)
Major, P.: On a non-parametric estimation of the regression function. Stud. Sci. Math. Hungar. **8**, 347–361 (1973)
Nadaraya, E.A.: On estimating regression. Theor. Probab. Appl. **9**, 141–142 (1964)
Nadaraya, E.A.: Remarks on nonparametric estimates for density functions and regression curves. Theor. Probab. Appl. **15**, 134–137 (1970)
Nadaraya, E.A.: Some limit theorems related to nonparametric estimates of regression curve (in Russian), Bull. Acad. Sci. Georgian S.S.R. **71**, 57–60 (1973)
Nadaraya, E.A.: The limit distribution of the quadratic deviation of nonparametric estimates of the regression function. Soobshch. Akad. Nauk. Gruz. SSR **74**, No. 1, 33–36 (1974) (In Russian)
Noda, K.: Estimation of a regression function by the Parzen kernel-type density estimators. Ann. Inst. Statist. Math. **28**, No. 2, 221–234 (1976)
Parzen, E.: On estimation of a probability density function and mode. Ann. Math. Statist. **33**, 1065–1076 (1962)
Reiss, R.-D.: Consistency of a certain class of empirical density functions. Metrika **22**, 189–203 (1975)
Révész, P.: On strong approximation of the multidimensional empirical process. Ann. Probab. **4**, 729–743 (1976)
Révész, P.: On the nonparametric estimation of the regression function. Problems Control. Inform. Theory **8**, 297–302 (1979)
Rosenblatt, M.: Remarks on a multivariate transformation. Ann. Math. Statist. **23**, 470–472 (1952)
Rosenblatt, M.: Remarks on some nonparametric estimators of a density function. Ann. Math. Statist. **27**, 832–837 (1956)
Rosenblatt, M.: Conditional probability density and regression estimates. In Multivariate Analysis II, ed. Krishnaiah, 25–31 (1969)
Rosenblatt, M.: On the maximal deviation of a k-dimensional density estimator. Ann. Probab. **4**, 1009–1015 (1976)
Royall, R.M.: A class of nonparametric estimators of a nonlinear regression function. Ph.D. dissertation, Stanford University (1966)
Schuster, E.F.: Joint asymptotic distribution of the estimated regression function at a finite number of distinct points. Ann. Math. Statist. **43**, 84–88 (1972)

Schuster, E.F., Yakowitz, S.: Contributions to the theory of nonparametric regression, with application to system identification. Ann. Statist. **7**, 139–149 (1979)
Silverman, B.W.: On a Gaussian process related to multivariate probability density estimation. Math. Proc. Cambridge Philos. Soc. **80**, 135–144 (1976)
Silverman, B.W.: Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. Ann. Statist. **6**, 177–184 (1978)
Stone, C.J.: Consistent nonparametric regression. Ann. Statist. **5**, 595–620 (1977)
Stute, W.: A law of logarithm for kernel density estimators. Ann. Probability **10**, 414–422 (1982)
Tusnády, G.: A remark on the approximation of the sample d.f. in the multidimensional case. Period. Math. Hungar. **8**, 53–55 (1977)
Wandl. H.: On kernel estimation of regression functions. Wiss. Sit. z. Stoch. WSS-03, 1–25 (1980)
Watson, G.S.: Smooth regression analysis. Sankhyā Ser. A **26**, 359–372 (1964)

Received January 31, 1981; in revised form March 26, 1982