Serdica

Bulgariacae mathematicae publicationes

Сердика

Българско математическо списание

NONPARAMETRIC ESTIMATION OF CURVES

WALTER SCHLEE

In this paper we deal with nonparametric estimators for multivariate distribution functions, multivariate densities, multivariate regression curves and derivatives hereof. Consistency and asymptotic normality are proved. Optimal parameters of the estimators as functions of the sample size are computed. Estimation and test of the structures of the theoretical curves are discussed.

The given estimators are some kind of kernel estimates, which are investigated for instance by Parzen (1962), Murthy (1965), Nadaraya (1965), Cacoullos (1966) and Bhattacharya (1967) in the case of a density. In the case of a regression curve there are papers by Nadaraya (1964) and (1965), Schuster (1972), Stone (1977) and Meyer (1977). The paper of Tukey (1961) gives a regressogram, which is an analogue to the histogram in the case of regression.

The paper of Stone (1977) concerns only the consistency of nonparametric regression by weighted sums. However our aim is to give the actual bias and limiting expectations, variances and covariances. In our cases the consistency follows immediately hereof. Meyer (1977) uses splines for deriving convergence bounds like the ones described below. Smoothed distribution functions are given e. g. by Penkov (1976) and Wahba (1976). For distribution functions we use splines, which are piecewise polynomial of at least third order.

The main emphasis of this paper lies on the fact that there are investigated the statistical properties of the really computed regression curve resp. density resp. distrtbution function estimator (besides numerical contamination caused by summation and miltiplication). Indeed no further discretization is needed here in contrary to the usual kernel estimates and the functional series expansions. This is achieved by some simple histospline whose complexity, however increases with the needed number of derivatives. A further advantage is the possibility of on-line computation of the estimators.

1. Description of the estimators. In the following we consider random variables (X, Y), whereas X is a q-dimensional r. v. and Y a q-dimensional r. v. Developed are estimators of the distribution function F and density f of X and their derivatives and an estimator of the regression of Y on X, that is $g(x) = \mathscr{E}(y|x)$ and the derivatives of the product $g(x)f(x)$. Let us have a sample (X_i, Y_i) of size n and let $r=(r_1, \ldots, r_q)$ denote that the r_j-th derivative in the variable x_j, $j=1, 2, \ldots q$ is taken. By $r_j = -1$ we denote the integral with respect to the variable x_j. Let be $|r| = \Sigma_{j=1}^q r_j$. Sometimes r denotes an index vector with only one component, but confusion is avoided. The estimator for the density resp. d. f. ($r=(-1, -1, \ldots, -1)$) and their derivatives is proposed as follows:

$$(1.1) \qquad f_n^{(r)}(x_1, \ldots, x_q) = h^{-q-|r|} \sum_{i_1, \ldots, i_q} \prod_{j=1}^{q} k^{(r_j)}\left(\frac{x_j - i_j\beta}{h}\right) \Delta(i) F_n,$$

($f_n^{(1, \ldots, -1)}$ is the smoothed estimator of the d. f.). The estimator for the product $f(x)g(x)$ and its derivatives is

SERDICA Bulgaricae mathematicae publicationes. Vol. 5, 1979, p. 186—203.

$$(1.2) \qquad \gamma_n^{(r)}(x_1, \ldots, x_q) = h^{-q-|r|} \sum_{i_1, \ldots, i_q} y_{(i)} \prod_{j=1}^{q} k^{(r_j)}\left(\frac{x_j - i_j\beta}{h}\right) \Delta(i) F_n.$$

n (1.1) and (1.2) and in the sequel the sum is taken over all integers $i_j = -\infty, \ldots, +\infty$, $j = 1, 2, \ldots, q$ unless there is explicitly given a restriction. A suitable estimator for the regression is

$$(1.3) \qquad g_n(x_1, \ldots, x_q) = \begin{cases} \gamma_n(x_1, \ldots, x_q)/f_n(x_1, \ldots, x_q) & \text{if } f_n(x) > 0, \\ 0 & \text{if } f_n(x) = 0. \end{cases}$$

The explanation of the parameters is as follows: $i = (i_1, \ldots, i_q)$ denotes an index vector. There is given a partition of the $\mathbb{R}^q$ into intervals with side length $\beta(h)$ and h, β are parameters ≥ 0. They are connected by thé following limits:

$$(1.4) \qquad \lim_{n\to\infty} h(n) = 0, \quad \lim_{h\to 0} \beta(h)/h = 0, \quad \lim_{n\to\infty} nh^q(n) = \infty.$$

Occasionally additional limits are demanded. By $\overline{y}_{(i)}$ we denote the arithmetic mean of all sample variables y, whose corresponding x lies in the interval $\Pi_{l=1}^q](i_l-1)\beta, i_l\beta]$; further F_n is the empirical distribution function of X and $\Delta(i)u$ denotes the q-folded difference of a function $u: \mathbb{R}^q \to \mathbb{R}$ with respect to the chosen interval partition defined by the side length β:

$$(1.5) \qquad \Delta(i)u = \sum_{j_1, \ldots, j_q} u((i_1-j_1)\beta, \ldots, (i_q-j_q)\beta)(-1)^{\Sigma j_k},$$

whereas the summation is taken over all values $j_1, j_2, \ldots, j_q = 0,1$. Therefore $\Delta(i)F_n$ resp. $\Delta(i)F$ is the empirical resp. theoretical probability that X lies in the interval $\Pi_{l=1}^q](i_l-1)\beta, i_l\beta]$.

It is assumed that the function $k: \mathbb{R} \to \mathbb{R}$ has the properties:

$$(1.6.1) \qquad k(x) \geq 0, \quad k(x) \equiv 0 \ \forall\, x: |x| > C_x,$$

(1.6.2) k is Riemann-integrable,

$$(1.6.3) \qquad \int_{\mathbb{R}} k(x)dx = 1,$$

$$(1.6.4) \qquad \sup_x |k(x)| \leq C_k.$$

If the r_j-th derivatives are considered there will be assumed that

(1.6.5) k is continuously differentiable up to the (r_j+1)-th order.

Generalizations are straightforward like replacing

(i) h^{-q} by $\Pi_{j=1}^q h_j^{-1}$,

(ii) $\Pi_{j=1}^q k((x_i - i_j\beta)/h)$ by $k((x_1 - i_1\beta_1)/h_1, \ldots\ldots\ldots, (x_q - i_q\beta_q)/h_q)$,

(iii) (1.6.1) by $k(x) \sim O(x^{-1})$ for $x \to \pm\infty$,

(iv) or dropping the positivity of k.

Only the case $\Pi_{j=1}^q k((x_j - i_j\beta_j)/h_j)$ with different β_j, h_j is practically interesting. In order to keep small the number of indices we dispense with the explicit elaboration of this case. It will be seen that all results remain valid. We do not take into account the remaining generalizations because there seems to be no real application.

The announced immediate computability of the above estimators is realized only if we choose an appropriate function k. But we would like to confirm that the below described special form is not used in the proofs. This special form is only proposed for computational convenience.

Constructing k we begin with the $(r+1)$-th derivative, if the r-th is needed. The derivative $k^{(r)}$ is defined as a linear spline on 2^{2+r} intervals with equal length. For simplicity let the length be equal to 1. The lefthand endpoint is assumed to be $a-2^{r+1}$ and the righthand endpoint $a+2^{r+1}$ with integer a. The maxima resp. minima are ε resp. $-\varepsilon$, whereas ε is given by the condition (1.6.3). We assume that k is symmetric with respect to a and non-negative and takes its maximum at the point a. Further it is assumed $k^{(i)}(a-2^{r+1})=0$, $k^{(i)}(a+2^{r+1})=0$, $i=0, 1, 2, \ldots, r$.

These splines may be defined recursively beginning with a spline $k(\cdot, a, 0)$ of order 0 with the following second derivative:

$$k''(x, a, 0)=\begin{cases} \varepsilon_0 & \text{if } a-2\leq x<a-1, \\ -\varepsilon_0 & \text{if } a-1\leq x<a, \\ -\varepsilon_0 & \text{if } a\leq x<a+1, \\ \varepsilon_0 & \text{if } a+1\leq x\leq a+2. \end{cases}$$

This implies

$$k'(x, a, 0)=\begin{cases} \varepsilon_0(x-(a-2)) & \text{if } a-2\leq x\leq a-1, \\ -\varepsilon_0(x-a) & \text{if } a-1\leq x\leq a, \\ -\varepsilon_0(x-a) & \text{if } a\leq x\leq a+1, \\ \varepsilon_0(x-(a+2)) & \text{if } a+1\leq x\leq a+2; \end{cases}$$

$$(1.7)\qquad \frac{k(x, a, 0)}{\varepsilon_0}=\begin{cases} (x-a)^2/2+2(x-a)+2 & \text{if } a-2\leq x\leq a-1, \\ -(x-a)^2/2+1 & \text{if } a-1\leq x\leq a, \\ -(x-a)^2/2+1 & \text{if } a\leq x\leq a+1, \\ (x-a)^2/2-2(x-a)+2 & \text{if } a+1\leq x\leq a+2; \end{cases}$$

$$(1.8)\qquad \frac{k^{(-1)}(x, a, 0)}{\varepsilon_0}=\begin{cases} 0 & \text{if } x\leq a-2, \\ (x-a)^3/6+(x-a)^2+2(x-a)+4/3 & \text{if } a-2\leq x\leq a-1, \\ -(x-a)^3/6+(x-a)+1 & \text{if } a-1\leq x\leq a, \\ -(x-a)^3/6+(x-a)+1 & \text{if } a\leq x\leq a+1, \\ (x-a)^3/6-(x-a)^2+2(x-a)+2/3 & \text{if } a+1\leq x\leq a+2, \\ 2 & \text{if } a+2\leq x. \end{cases}$$

The recursive definition is:

$$(1.9)\qquad \frac{k'(x, a, i)}{\varepsilon_i}=\begin{cases} k(x, a-2^i, i-1)/\varepsilon_{i-1} & \text{if } a-2^{i+1}\leq x\leq a, \\ -k(x, a+2^i, i-1)/\varepsilon_{i-1} & \text{if } a\leq x\leq a+2^{i+1}; \end{cases}$$

and the symmetry yields $\varepsilon_i=2^{-i-1}$.

The algorithmic computation of the spline of order i begins with its $a+(2)$-th derivative, which has the values $\pm\varepsilon_i$ only. The symmetry described by (1.9) yields the following table of the signs of the $(i+2)$-th derivatives:

	a
$k^{(2)}(\cdot, a, 0)$:	$+ - \vdots - +$
$k^{(3)}(\cdot, a, 1)$:	$+ - - + \vdots - + + -$
$k^{(4)}(\cdot, a, 2)$:	$+ - - + - + + - \vdots - + + - + - - +$.

The table can be continued easily.

By integration we get the lower derivatives of k or k itself or its integral. Let the l-th derivative in the interval $(j, j+1)$ be denoted by

$$f_j^l(x)=\sum_{k=0}^{i+2-1} c_{jk}^l x^k. \tag{1.10}$$

The equality $f_j^{l-1}(x)=f_{j-1}^l(j)+\sum_{k=0}^{i+2-1} c_{jk}^1(x^{k+1}-j^{k+1})/k$ implies the recursion formulas

$$\begin{aligned} &c_{j0}^{i+2}=\pm\varepsilon_i \quad \text{according to the above table,}\\ &c_{jk}^{l-1}=c_{jk-1}^l/k, \quad k=1, 2, \ldots, (i+2-1),\\ &c_{j0}^{l-1}=f_{j-1}^{l-1}(j)-\sum_{k=1}^{i+2-1} c_{jk}^{l-1} j^k,\\ &l=i+1,\ i,\ i-1, \ldots \end{aligned} \tag{1.11}$$

Computational forms of the estimators are

$$\begin{aligned} f_n^{(r)}(x)&=(nh^{q+|r|})^{-1}\sum_{l=1}^n \prod_{j=1}^q k^{(r_j)}\left(\frac{x_j-i_j(u_l)\beta}{h}\right)\\ \gamma_n^{(r)}(x)&=(nh^{q+|r|})^{-1}\sum_{l=1}^n y_l \prod_{j=1}^q k^{(r_j)}\left(\frac{x_j-i_j(u_l)\beta}{h}\right), \end{aligned} \tag{1.12}$$

whereas $i_j(u_l)$ denotes the index $i=(i_1, \ldots, i_q)$ of the sample value (y_l, u_l) with $u_l\in\prod_{s=1}^q](i_s-1)\beta, i_s\beta]$. The formulas (1.12) show that the estimators can be computed on-line adding step by step $\prod_{j=1}^q k^{(r_j)}((x_j-i_j(u_l)\beta)/h)$ resp. $y_l\prod_{j=1}^q k^{(r_j)}((x_j-i_j(u_l)\beta)/(h)$. In the sequel we use the abbreviation $\widehat{k}^{(r)}(x, i)=\prod_{j=1}^q k^{(r_j)}((x_j-i_j\beta)/h)$ and ζ_l, η_l denote appropriate values of the interval $\prod_{j=1}^q](i_j-1)\beta, i_j\beta]$.

2. Some basic lemmas. In the sequel the properties (1.6) are assumed valid. The first three lemmas are needed in order to handle the derivatives and the integrals.

L e m m a 1. *Let v_j: $\mathbb{R}\to\mathbb{R}$, $j=1, \ldots, q$ be of bounded variation and u $\mathbb{R}^q\to\mathbb{R}$ a bounded function, $\widehat{v}=\prod_{j=1}^q v_j$. If $\Delta(i)\widehat{v}$ is zero outside a bounded domain, then*

$$\sum_{i_1, \ldots, i_q}\widehat{v}(i\beta)\,\Delta(i)u=(-1)^q\sum_{i_1, \ldots, i_q} u(i\beta)\Delta(i+1)\widehat{v} \tag{2.1}$$

$$=(-1)^q \sum_{i_1,\dots,i_q} u(i\beta) \prod_{j=1}^{q} (v_j((i_j+1)\beta)-v_j(i_j\beta)).$$

The proof follows from the fact that the righthandside of these equalities is absolutely convergent for a given interval partition.

Lemma 2. *Let $u: \mathrm{R} \to \mathrm{R}$ be $(r+1)$-times continuously differentiable, then*

$$(2.2)\quad u^{(r)}(x-i\beta)=\beta^{-r} \sum_{l=0}^{r} \binom{r}{l} u(x-(i-l)\beta)(-1)^{r-1}-\beta C_{r+1} u^{(r+1)}(x-i\beta+\eta_i^{r+1}).$$

C_{r+1} is a constant which depends on u and r but does not depend on β.

Proof. The proof follows by induction. For $r=1$ the formula (2.2) is implied by the power series of u. For $u^{(r)}$ the power series yield

$$u^{(r+1)}(x-i\beta)=\beta^{-1}(u^{(r)}(x-(i-1)\beta)-u^{(r)}(x-i\beta))-C_2 u^{(r+1)}(x-i\beta+\eta_i^{r+2}).$$

Assuming now that (2.2) is valid for r the formula (2.2) may be proved for $r+1$.

Lemma 3. *Let $u: \mathrm{R}^q \to \mathrm{R}$ $(r+1)$-times continuously differentiable and $u^{(r+1)}$ bounded and Riemann-integrable, then*

$$(2.3)\quad u^{(r)}(x_1-i_1\beta,\dots,x_q-i_q\beta)$$

$$=\beta^{-|r|} \sum_{l_1=0}^{r_1} \dots \sum_{l_q=0}^{r_q} \prod_{j=1}^{q} \binom{r_j}{l_j} u(x-(i-1)\beta)(-1)^{|r|-|l|}-\beta \tilde{u}(x-i\beta+\eta_i^{r+1}),$$

where $\tilde{u}$ is a bounded Riemann-integrable function.

The proof follows by lemma 2.

Lemma 4. *For all $l\geq 1$, $r_j\geq 0$, $j=1,2,\dots,q$*

$$(2.4)\quad \lim_{h\to 0} \sum_{i_1,\dots,i_q} (\widehat{k}^{(r)}(x,i))^l \left(\frac{\beta}{h}\right)^q = \prod_{j=1}^{q} \int (k^{(r_j)}(t))^l dt.$$

Proof. The formal identity

$$\sum_{i_1,\dots,i_q} \prod_{j=1}^{q} \left(k^{(r_j)}\left(\frac{x_j-i_j\beta}{h}\right)\right)^l \left(\frac{\beta}{h}\right)^q = \prod_{j=1}^{q} \left[\sum_{i=-\infty}^{\infty} \left(k^{(r_j)}\left(\frac{x_j-i\beta}{h}\right)\right)^l \frac{\beta}{h}\right]$$

allows a reduction to the 1-dimensional case. $\sum_{i=-\infty}^{\infty}$ can be replaced by $\sum_{i:|(x-i\beta)/h|\leq C_x}$, because $k^{(r)}$ is outside the domain described by C_x equal zero. This sum has only finitely many summands and is a Riemann sum which converges to the corresponding integral.

Lemma 5. *Let $r_j\geq 0$ and $u: \mathrm{R}^q \to \mathrm{R}$ be $(r_1+1,\dots,r_q+1)$-times continuously differentiable in a neighbourhood of x. Further let*

$$\sup_x |u^{(v)}(x)|\leq c_u^v,\ \xi(i)\in \prod_{j=1}^{q}](i_j-1)\beta,\ i_j\beta],\quad \lim_{h\to 0} \beta(h)h^{-|r|-1}=0.$$

Then

(2.5) $$\lim_{h\to\infty} h^{-|r|} \sum_{i_1,\ldots,i_q} \widehat{k}^{(r)}(x,i)u(\xi(i))\left(\frac{\beta}{h}\right)^q = u^{(r)}(x)(\int k(t)dt)^q.$$

In the case of $|r|=0$ *only the continuity of* u *in a neighbourhood of* x *is needed. It is not necessary that* (1.6.3) *is fulfilled.*

Proof. In the case $|r|=0$:

$$\left|\sum_i \widehat{k}(x,i)\left(\frac{\beta}{h}\right)^q u(\xi(i)) - u(x)(\int k(t)\,dt)^q\right| \leq \left|\sum_{i:|x_j-i_j\beta|\leq\delta} \widehat{k}(x,i)\left(\frac{\beta}{h}\right)^q (u(\xi_i)-u(x))\right|$$

$$+\left|\sum_{i:|x_j-i_j\beta|>\delta} \widehat{k}(x,i)\left(\frac{\beta}{h}\right)^q (u(\xi_i)-u(x))\right| + \left|\sum_i \widehat{k}(x,i)\left(\frac{\beta}{h}\right)^q - (\int k(t)dt)^q\right| |u(x)|$$

$$= \sup_{y:|x-y|\leq\delta+\beta} |u(x)-u(y)| \sum_{i:|x_j-i_j\beta|\leq\delta} \widehat{k}(x,i)\left(\frac{\beta}{h}\right)^q$$

$$+2\sup_x u(x) \left|\sum_{i:|x_j-i_j\beta|>\delta} \widehat{k}(x,i)\left(\frac{\beta}{h}\right)^q + \sup_x u(x)\right| \left|\sum_i \widehat{k}(x,i)\left(\frac{\beta}{h}\right)^q - (\int k(t)dt)^q\right|.$$

The continuity of u at the point x implies the existence of a δ and a $\beta(h)$ sufficiently small so that $\sup\{|u(x)-u(y)|<\eta : y: |x-y|\leq\delta+\beta\}$ for every $\eta>0$. Now we choose an h (and $\beta(h)$ eventually decreases) with $\delta/h>C_x$. Then the second sum is equal to zero.

Lemma 4 implies that $\sum_{i:|x_j-i_j\beta|\leq\delta} \widehat{k}(x,i)(\beta/h)^q$ is bounded and that the third summand converges to zero.

In the case $|r|>0$: According to lemma 2

$$\widehat{k}^{(r)}(x,i) = \left(\frac{\beta}{h}\right)^{|r|} \sum_{l_1,\ldots,l_q=0}^{r_1,\ldots,r_q} \prod_{j=1}^q \binom{r_j}{l_j} \widehat{k}(x,i-l)(-1)^{|r|-|l|} - (\beta/h)\widetilde{\widehat{k}}(x-i\beta)+\eta_i^{r+1}),$$

$$h^{|r|} \sum_i \widehat{k}^{(r)}(x,i)\, u(\xi_i)\left(\frac{\beta}{h}\right)^q$$

$$= \sum_i\sum_l \prod_{j=1}^q \binom{r_j}{l_j} \widehat{k}(x,i-1)(-1)^{|r|-|l|} u(\xi_i)\left(\frac{\beta}{h}\right)^q \beta^{-|r|}$$

$$- h^{-|r|}(\beta/h) \sum_i \widetilde{\widehat{k}}(x-i\beta+\eta_i^{r+1})\, u(\xi_i)\left(\frac{\beta}{h}\right)^q.$$

With the propositions about k and about $h^{-|r|-1}\beta$ the second summand converges to zero. Using $i_j := i_j - l_j$ and lemma 3 the first summand is transformed to

$$\beta^{-r} \sum_i\sum_l \prod_{j=1}^q \binom{r_j}{l_j} \widehat{k}(x,i)(-1)^{|r|-l} u(\xi_{i+1})\left(\frac{\beta}{h}\right)^q$$

$$= \sum_i \left(\frac{\beta}{h}\right)^q \widehat{k}(x,i)(u^{(r)}(\xi_i)+\beta\widetilde{u}(\xi_i+\eta_i^{r+1})).$$

Lemma 2 and the case $|r|=0$ imply that this quantity converges to the one of (2.5).

Lemma 6. *For all $u, v>0$ and $(x_1, \ldots, x_q)\neq(y_1, \ldots, y_q)$ and bounded g it holds*

$$\lim_{h\to 0} \sum_{i_1, \ldots, i_q} (\widehat{k}^{(r_1)}(x, i))^u(\widehat{k}^{(r_2)}(y, i))^v \left(\frac{\beta}{h}\right)^q g(\xi_i) = 0. \tag{2.6}$$

Proof.

$$\left|\sum_i (\widehat{k}^{(r_1)}(x, i))^u(\widehat{k}^{(r_2)}(y, i))^v \left(\frac{\beta}{h}\right)^q g(\xi_i)\right|$$

$$\leq \sup_x |g(x)| \sup_x |\widehat{k}^{(r_1)}(x)|^u \sum_{i:\,|(x_j-i_j\beta)/h|\leq\delta/h} (\widehat{k}^{(r_2)}(y, i))^v \left(\frac{\beta}{h}\right)^q$$

$$+ \sup_x |g(x)| \sup_y |\widehat{k}^{(r_2)}(y)|^v \sum_{i:\,|(x_j-i_j\beta)/h|>\delta/h} (\widehat{k}^{(r_1)}(x, i))^u \left(\frac{\beta}{h}\right)^q.$$

At first we choose h small enough so that $\delta/h>C_x$. Then the second summand is zero. Further we decrease h so that $|x_j-y_j|>2C_xh$. Then the first summand is zero.

3. Bias and consistency. Estimating $f^{(r)}$, $(gf)^{(r)}$, $|r|>0$, it is possible to give a limiting variance in terms of f resp. (gf).

The consistency follows if we additionally suppose

(i) $\lim_{k\to\infty} nh^{q+2|r|}=\infty$

(ii) that the bias converges to 0.

These are restrictions to both $\beta(h)$ and $h(n)$.

Theorem 1. *Let $r_j\geq 0$ and the density f be $(r+1)$-times continuously differentiable. Then*

$$\lim_{n\to\infty} \mathcal{E}f_n^{(r)}(x)=f^{(r)}(x), \tag{3.1.1}$$

$$\lim_{n\to\infty} \operatorname{var}((nh^q)^{1/2} h^{|r|} f_n^{(r)}(x))=f(x) \prod_{j=1}^{q} \int (k^{(r_j)}(t))^2\,dt. \tag{3.1.2}$$

If f is $(r+2)$-times continuously differentiable, k symmetric to zero and $C_{i,n}$, $i=1, 2, 3, r$ are suitable constants then

$$b[f_n^{(r)}(x)]=\mathcal{E}f_n^{(r)}(x)-f^{(r)}(x)=\frac{h^2}{2}\left[\sum_{j=1}^{q} \frac{\partial^2 f^{(r)}}{\partial x_i^2}\bigg|_x \left(\int t^2k(t)\,dt\right)+C_{3,n}\cdot\frac{\beta}{h}\right] \tag{3.1.3}$$

$$+\left(\frac{\beta}{h}\right)(C_{2,n}+h^{-|r|}C_{r,n})+\beta C_{1,n}.$$

If $|r|=0$ then $C_{r,n}=0$.

Proof. Proof of (3.1.1): $\mathcal{E}f_n^{(r)}=h^{-|r|}\sum_i\widehat{k}^{(r)}(x, i)(\beta/h)^q\Delta(i)F/\beta^q$, $\Delta(i)F/\beta^q\approx f(\xi_i)$ and lemma 5 imply the assertion.

Proof of (3.1.2):

$$\operatorname{Var}((nh^q)^{1/2} h^{|r|} f_n^{(r)}(x))=nh^{-q}\sum_i\sum_j \widehat{k}^{(r)}(x, i)\widehat{k}^{(r)}(x, j)\mathcal{E}(\Delta(i)(F_n-F)\Delta(j)(F_n-F)).$$

If $i\neq j$: $\mathcal{E}(\Delta(i)(F_n-F)\Delta(j)(F_n-F))=-\Delta(i)F\cdot\Delta(j)F/n$. If $i=j$: $\mathcal{E}(\Delta(i)(F_n-F)^2$

$=\Delta(i)F(1-\Delta(i)F)/n$. Therefore, the summands with $i\neq j$ are of order h^{2q} and are added up to order h^q according to lemma 5. Consequently this sum converges to zero. The remaining sum yields the stated value.

Proof of (3.1.3). According to lemma 3 and the proof of lemma 5 there exist bounded quantities $C_{1,n}$, $C_{r,n}$ with

$$b[f_n^{(r)}]=h^{-|r|}\sum_i \widehat{k}^{(r)}(x,i)\left(\frac{\beta}{h}\right)^q f(\xi_i)-f^{(r)}(x)$$

$$=\sum_i \widehat{k}(x,i)\left(\frac{\beta}{h}\right)^q (f^{(r)}(\xi_i)-f^{(r)}(x))+\left(\sum_i \widehat{k}(x,i)\left(\frac{\beta}{h}\right)^q-1\right)f^{(r)}(x)$$

$$+h^{-|r|}(\beta/h)C_{r,n}+\beta C_{1,n},$$

taking into account that $(\Sigma_i \widehat{k}(x,i)(\beta/h)^q-1)$ converges to zero like $(\beta/h)C_2$ with a suitable constant C_2. The Taylor series of $f^{(r)}$ yields

$$h\left(\sum_{j=1}^{q}\sum_i \widehat{k}(x,i)\frac{\Delta_{ji}}{h}(\beta/h)^q\frac{\partial f^{(r)}}{\partial x_j}\bigg|_{\xi_i}\right)+\frac{h^2}{2}\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_i \widehat{k}(x,i)\frac{\Delta_{ji}}{h}\frac{\Delta_{li}}{h}\left(\frac{\beta}{h}\right)^q\frac{\partial^2 f^{(r)}}{\partial x_j\partial x_l}\bigg|_{\xi_i}$$

$$+(\beta/h)C_{2,n}+h^{-|r|}(\beta/h)C_{r,n}+\beta C_{1,n}.$$

The symmetry of k implies the first sum to be 0 and the second sum reduces to a sum with $l=j$. Therefore, the assertion is true.

Theorem 2. *Let $\widetilde{F}_n=f_n^{(-1,\dots,-1)}$ be the smoothed d. f. estimate. If k satisfies the conditions* (1.6.1)—(1.6.4) *then*:

$$\lim_{n\to\infty}\mathcal{E}\widetilde{F}_n(x)=F(x), \tag{3.2.1}$$

$$\lim_{n\to\infty}\mathrm{Var}(\sqrt{n}\,\widetilde{F}_n(x))=F(x)(1-F(x)). \tag{3.2.2}$$

If k and F are continuously differentiable in a neighbourhood of x and k is symmetric to zero then the bias is

$$b[\widetilde{F}_n(x)]=\mathcal{E}\widetilde{F}_n(x)-F(x)=C_{1,n}\frac{\beta}{h}+\frac{h^2}{2}\left[\sum_{i=1}^{q}\frac{\partial^2 F}{\partial x_i^2}\bigg|_x\left(\int t^2k(t)dt\right)+C_{2,n}\frac{\beta}{h}\right]. \tag{3.2.3}$$

Proof. Proof of (3.2.1): Using lemma 1 and $x-i\beta\leq\eta_i h\leq x-(i+1)\beta$ we get

$$\mathcal{E}(\sum_i \widehat{k}^{(-1)}(x,i)\Delta(i)F_n)=(-1)^q\sum_i F(i\beta)\Delta(i+1)\widehat{k}^{(-1)}=(-1)^{2q}\sum_i F(i\beta)\,\widehat{k}(\eta_i,i)(\beta/h)^q.$$

This quantity converges to $F(x)$ according to lemma 5.

Proof of (3.2.2): Taking into account that

$$\mathcal{E}(n\Delta(i)(F_n-F)\Delta(j)(F_n-F))=\begin{cases}-\Delta(i)F(\Delta(j)F & \text{if } i\neq j,\\ \Delta(i)F\ (1-\Delta(i)F) & \text{if } i=j,\end{cases}$$

we have

$$\mathcal{E}(n\sum_i\sum_j \widehat{k}^{(-1)}(x,i)\,\widehat{k}^{(-1)}(x,j)\ \Delta(i)(F_n-F)\ \Delta(j)(F_n-F)$$

$$= -\sum_i \sum_j \widehat{k}^{(-1)}(x,i)\widehat{k}^{(-1)}(x,j)\Delta(i)F\Delta(j)F + \sum_i (\widehat{k}^{(-1)}(x,i))^2 \Delta(i)F.$$

According to lemma 1 the second sum can be transformed to

$$\sum_i F(i\beta)\Delta(i+1)\,(\widehat{k}^{(-1)}(x,\,i))^2 = \sum_i F(i\beta)((\widehat{k}^{(-1)}(\eta_i,\,i))^2)^{(1,\,\ldots,\,1)}(\beta/h)^q.$$

Lemma 4 implies the convergence of this sum to

$$F(x)\prod_{j=1}^{q}\int d\,(k^{(-1)})^2 = F(x).$$

Proof of (3.2.3). Considering the proof of (3.2.1) we get the following equality:

$$\sum_i F(i\beta)\widehat{k}(\eta_i,i)\,(\beta/h)^q - F(x)$$

$$= \sum_i F(i\beta)(\widehat{k}(\eta_i,\;i) - \widehat{k}(x,i))\,(\beta/h)^q + \sum_{i\,:\,|x-i\beta|\leq\delta}(F(i\beta)-F(x))\widehat{k}(x,\,i)\,(\beta/h)^q$$

$$+ \sum_{i\,:\,|x-i\beta|>\delta}(F(i\beta)-F(x))\widehat{k}(x,\,i)\left(\frac{\beta}{h}\right)^q + \Big(\sum_i \widehat{k}(x,\,i)\left(\frac{\beta}{h}\right)^q - 1\Big)F(x).$$

$\delta>0$ is arbitrary and we may consider only such h's which are small enough to yield $\delta/h>C_x$. Using the Taylor series of $\widehat{k}$ and F and the symmetry of k we obtian the assertion analoguously to the above proofs.

Now $\mathfrak{X}$ denotes the sample vector $(X_1,\ldots,X_n)$.

Theorem 3. *There is assumed that bounded versions of the conditional expectation $\mathcal{E}(Y|X)$ and the conditional covariance matrix exist. In the sequel we denote $\mathcal{E}(Y|X)$ by $g(x)$ and the conditional covariance matrix* $(\mathrm{Cov}\,(Y_\nu, Y_\mu|X))$ *by* $(\sigma_{\nu\mu}(x))$ *and σ^2 denotes the vector* $(\sigma_1^2,\ldots,\sigma_q^2)=(\sigma_{11},\ldots,\sigma_{qq})$.

The product $g(x)f(x)$ is supposed to be $(r+1)$-times continuously differentiable, then

$$\lim_{n\to\infty}\mathcal{E}\gamma_n^{(r)}(x) = (g(x)f(x))^{(r)}, \tag{3.3.1}$$

$$\lim_{n\to\infty}\mathrm{Var}\,((nh^q)^{1/2}\,h^{|r|}\,\gamma_n^{(r)}(x)) = \sigma^2(x)f(x)\prod_{j=1}^{q}\int(k^{(r_j)}(t))^2dt. \tag{3.3.2}$$

If $\nu\neq\mu$ then

$$\lim_{n\to\infty}\mathrm{Cov}\,((nh^q)^{1/2}h^{|r|}\gamma_{n\nu}^{(r)}(x),\;(nh^q)^{1/2}\,h^{|r|}\gamma_{n\mu}^{(r)}(x)) \tag{3.3.3}$$

$$= (\sigma_{\nu\mu}(x)+g_\nu(x)g_\mu(x)\,f(x)\prod_{j=1}^{q}\int(k^{(r_j)}(t))^2dt.$$

If gf is $(r+2)$-times continuously differentiable and k is symmetric with respect to zero then

$$b\,[\gamma_n^{(r)}(x)] = \frac{h^2}{2}\left[\sum_{j=1}^{q}\frac{\partial^2(fg)^{(r)}}{\partial x_j^2}\left(\int t^2k\,(t)dt\right)+C_{3,n}\,(\beta/h)\right] \tag{3.3.4}$$

$$+(\beta/h)\,(C_{2,n}+h^{-|r|}C_{r,n})+\beta C_{1,n}$$

with $C_{r,n}=0$ if $|r|=0$.

P r o o f. Proof of (3.3.1):

$$\mathcal{E}\gamma_n^{(r)}(x)=\mathcal{E}(\mathcal{E}(\gamma_n^{(r)}(x)\mid\mathfrak{X}))=h^{-|r|}\sum_i \widehat{k}^{(r)}(x,i)\,g(\xi_i)\left(\frac{\beta}{h}\right)^q\frac{\Delta(i)F}{\beta^q}.$$

The proof is similar to the proofs of the above theorems.

Proof of (3. 3. 3) and (3. 3. 2):

$$\mathrm{Cov}\,((nh^q)^{1/2}h^{|r|}\gamma_{n\nu}^{(r)}(x),\ (nh^q)^{1/2}h^{|r|}\gamma_{n\mu}^{(r)}(x))$$

$$=nh^{-q}\sum_i\sum_j \widehat{k}^{(r)}(x,i)\,\widehat{k}^{(r)}(x,j)\,\mathcal{E}(\bar{y}_{(i)\nu}\Delta(i)F_n-g_\nu(\xi_i)\Delta(i)F)\,.\,(\bar{y}_{(j)\mu}\Delta(j)F_n-g_\mu(\xi_i)\Delta(j)F).$$

The stochastic independence of $\bar{y}_{(i)}$ reduces this sum to $i_l=j_l,\ l=1,\dots,q$. The assertions are obtained by

$$\mathcal{E}(\bar{y}_{(i)\nu}\Delta(i)F_n-g_\nu(\xi_i)\Delta(i)F)\,.\,(\bar{y}_{(i)\mu}\Delta(i)F_n-g_\mu(\xi_i)\Delta(i)F)$$

$$=\mathcal{E}[(\bar{y}_{(i)\nu}-g_\nu(\xi_i))(\bar{y}_{(i)\mu}-g_\mu(\xi_i))(\Delta(i)F_n)^2+g_\nu(\xi_i)g_\mu(\xi_i)(\Delta(i)F_n-\Delta(i)F)^2$$

$$-(\bar{y}_{(i)\nu}-g_\nu(\xi_i))\Delta(i)F_ng_\mu(\xi_i)(\Delta(i)F_n-\Delta(i)F)-(\bar{y}_{(i)\mu}-g_\mu(\xi_i))\Delta(i)F_ng_\nu(\xi_i)(\Delta(i)F_n$$

$$-\Delta(i)F)]=(\sigma_{\nu\mu}(\eta_i)/n)\cdot\mathcal{E}\frac{(\Delta(i)F_n)^2}{\Delta(i)F_n}+g_\nu(\xi_i)g_\mu(\xi_i)\Delta(i)F(1-\Delta(i)F)/n.$$

Proof of (3. 3. 4). With appropriate constants it holds:

$$b[\gamma_n^{(r)}(x)]=h^{-|r|}\sum_i \widehat{k}^{(r)}(x,i)g(\xi_i)f(\eta_i)\left(\frac{\beta}{h}\right)^q-(g(x)f(x))^{(r)}$$

$$=\sum_i \widehat{k}(x,i)\left(\frac{\beta}{h}\right)^q((g(\xi_i)f(\eta_i))^{(r)}-(g(x)f(x))^{(r)})$$

$$+(\sum_i \widehat{k}(x,i)\left(\frac{\beta}{h}\right)^q-1)(g(x)f(x))^{(r)}+C_{r,n}h^{-|r|}(\beta/h)+\beta C_{1,n}.$$

T h e o r e m 4. *Let f and g be $(r+1)$-times continuously differentiable. Then*

(3.4.1) $$\lim_{n\to\infty}\mathrm{Cov}\,((nh^q)^{1/2}h^{|r|}f_n^{(r)}(x),\ (nh^q)^{1/2}h^{|r|}\gamma_n^{(r)}(x))=g(x)f(x)\prod_{j=1}^{q}\int(k^{(r_j)}(t))^2dt.$$

P r o o f.

$$\mathcal{E}[nh^{-q}\sum_i\sum_j \widehat{k}^{(r)}(x,i)\,(\bar{y}_{(i)\mu}\Delta(i)F_n-g_\mu(\xi_i)\Delta(i)F)\widehat{k}^{(r)}(x,j)\,(\Delta(j)F_n-\Delta(j)F]$$

$$=nh^{-q}\sum_i\sum_j k^{(r)}(x,i)\widehat{k}^{(r)}(x,j)g_\mu(\xi_i)\mathcal{E}(\Delta(i)(F_n-F)(\Delta(j)(F_n-F)).$$

The assertions follow by argument similar to the above.

T h e o r e m 5. *Let* $B_n:=b[\gamma_n^{(r)}(x)]-b[f_n^{(r)}(x)]\,(gf)^{(r)}/f^{(r)}$. *It is assumed that*

(3.5.1.) $$\lim_{n\to\infty} nh^{q+|r|}=\infty \text{ and } \lim_{n\to\infty} nh^{q+|r|}B_n^2<\infty\ \ \forall x.$$

Then it holds

(3.5.2) $$b[\gamma_n^{(r)}(x)/f_n^{(r)}(x)]=\mathcal{E}(\gamma_n^{(r)}/f_n^{(r)})-(gf)^{(r)}/f^{(r)}=B_n(x)/f^{(r)}(x)+O(n^{-2}h^{-2q-2|r|}).$$

Proof. The assumptions of the theorem yield the convergence of $f_n^{(r)}$ to $f^{(r)}$ in probability. Let be $v=(gf)^{(r)}/f^{(r)}$.

$$b\left[\frac{\gamma_n^{(r)}}{f_n^{(r)}}\right]=\sum_{i=0}^{\infty}(f^{(r)})^{-i-1}\mathcal{E}[(\gamma_n^{(r)}-vf_n^{(r)})(f_n^{(r)}-f^{(r)})^i].$$

For $i\geqq 1$ we have

$$|\mathcal{E}(\gamma_n^{(r)}-vf_n^{(r)})(f_n^{(r)}-f^{(r)})^i|\leq(\mathcal{E}(\gamma_n^{(r)}-vf_n^{(r)})^2\times\mathcal{E}(f_n^{(r)}-f^{(r)})^{2i})^{1/2},$$

$$\mathcal{E}(\gamma_n^{(r)}-vf_n^{(r)})^2=\mathrm{Var}\,(\gamma_n^{(r)})+v^2\,\mathrm{Var}\,(f_n^{(r)})-2v\,\mathrm{Cov}\,(\gamma_n^{(r)},f_n^{(r)})+B_n^2$$

$$=O(n^{-1}h^{-q-2|r|})+O(B_n^2),$$

$$\mathcal{E}(f_n^{(r)}-f^{(r)})^{2i}=(b[f_n^{(r)}])^{2i}+\sum_{l=2}^{2i}\binom{2i}{l}\mathcal{E}(f_n^{(r)}-\mathcal{E}f_n^{(r)})^l(b[f_n^{(r)}])^{2i-l}.$$

$b[f_n^{(r)}]$ is of order β and $\mathcal{E}(f_n^{(r)}-\mathcal{E}(f_n^{(r)})^l$ of order $(nh^{q+2|r|})^{-1/2}$. This proves (3.5.2).

4. Asymptotic normality. In this part we show the asymptotic normality of the r. v. $(Z_n^T(x_1),\ldots,Z_n^T(x_l))$, whereas

(4.1) $$Z_n(x)=(nh^{q+2|r|})^{1/2}\begin{pmatrix}\gamma_n^{(r)}(x)-\mathcal{E}\gamma_n^{(r)}(x)\\ f_n^{(r)}(x)-\mathcal{E}f_n^{(r)}(x)\end{pmatrix},$$

(4.2) $$Z_n^*(x)=(nh^q)^{1/2}h^{|r|}\begin{pmatrix}\gamma_n^{(r)}(x)-(gf)^{(r)}(x)\\ f_n^{(r)}(x)-f^{(r)}(x)\end{pmatrix}$$

if

(4.3) $$\lim_{n\to\infty}(nh^{q+2|r|})\,.\max\,(b[f_n^{(r)}],\,b[\gamma_n^{(r)}])=0\quad\forall x.$$

Moreover the asymptotic normality of $\gamma_n^{(r)}/f_n^{(r)}$ follows by (4.3) and a theorem which one finds e. g. in Witting [18, p. 49]. Theorem 9 gives the asymptotic normality of $\widetilde{F}_n$. The following theorem gives an l-variate central limit theorem for further use which the author did not find in the literature.

Theorem 6. *Let X_{nk}, $k=1,2,\ldots,n$, be stochastically independent l-dimensional r. v. with expectation vector* 0 *and covariance matrix $\mathcal{C}^{(nk)}=(\sigma_{ij}^{(nk)})$; X_{ink} denotes the i-th component of the vector X_{nk}. If the matrix*

(4.4.1) $$\mathcal{C}=\lim_{n\to\infty}\sum_{k=1}^{n}\mathcal{C}^{(nk)}$$

is strictly positive definite and

(4.4.2) $$\lim_{n\to\infty}\sum_{k=1}^{n}\max_i\mathcal{E}\,|X_{ink}|^3=0,$$

then $\lim_n Y_n = \lim_n \Sigma_{k=1}^n X_{nk}$ *is l-dimensionally normally distributed with expectation vector* 0 *and covariance matrix* $\mathcal{C} = \lim_{n\to\infty} \Sigma_{k=1}^n \mathcal{C}^{(nk)}$.

Proof. According to a well known theorem that an l-dimensional r. v. is normally distributed iff every (non-zero) linear combination of its components is 1-dimensionally normally distributed. Therefore, we consider the random variable $U_{nk} = \Sigma_{i=1}^l \alpha_i X_{ink}$ with some vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_l)$. To show the normal convergence of $\Sigma_{k=1}^n U_{nk}$ we use the normal convergence criterion of Loève [6, p. 307]:

Let be F_{nk} the distribution function of U_{nk}. There has to be

$$g_n(\varepsilon) = \sum_{k=1}^n \int_{|u|\geq\varepsilon} U^2 dF_{nk} \to 0.$$

This criterion is fulfilled if

$$\sum_{k=1}^n \mathcal{E}|U_{nk}|^3 \to 0 \text{ and} \tag{4.5.1}$$

$$\lim_{n\to\infty} s_n^2 = \lim_{n\to\infty} \sum_{k=1}^n \mathrm{Var}(U_{nk}) \text{ exists and is greater than zero.} \tag{4.5.2}$$

With the Minkowsky inequality we reformulate these conditions in terms of X_{ink}:

$$\mathcal{E}|U_{nk}|^3 \leq \left(\sum_{i=1}^l |\alpha_i| (\mathcal{E}|X_{ink}|^3)^{1/3}\right)^3 \leqq \left(\sum_{i=1}^l |\alpha_i|\right)^3 \max_i \mathcal{E}|X_{ink}|^3.$$

Now we see that (4.4.2) implies (4.5.1). The relation $\mathrm{Var}(U_{nk}) = \alpha^T \mathcal{C}^{(nk)} \alpha$ together with (4.4.1) implies (4.5.2) and Y_n is proved as normally distributed in the limit.

Theorem 7. *Let* $x_1, x_2, \ldots, x_l \in \mathbb{R}^q$ *denote continuity points of* $f^{(r)}$ *and* $(fg)^{(r)}$ *with* $f(x_i) > 0$ $i = 1, 2, \ldots l$. *The covariance matrix* $\tilde{\mathcal{C}}(x_i)$ *of* $z_n(x_i)$ *is assumed to be strictly positive definite and continuous in a neighbourhood of* x_i. *Further it is supposed that a bounded version* $\mu_3(x)$ *($\bar{q}$-dimensional vector) of the conditional third absolute central moment of* Y *exists and is continuous in a neighbourhood of* x_i.

Then $(z_n(x_1), \ldots, z_n(x_l))$ *is asymptotically normal with expectation zero and covariance matrix* $\mathcal{C}^*$ *which is block diagonal with block matrices* $\tilde{\mathcal{C}}(x_l)$:

$$\tilde{\mathcal{C}}(x) = \left(\begin{array}{c|c} (\tilde{\sigma}_{\nu\mu}(x)) & \begin{matrix} g_1(x) \\ \vdots \\ g_{\bar{q}}(x) \end{matrix} \\ \hline g_1(x) \ldots g_{\bar{q}}(x) & 1 \end{array}\right) \times f(x) \prod_{j=1}^q \int (k^{(r_j)}(t))^2 dt, \tag{4.6}$$

$$\tilde{\sigma}_{\nu\nu}(x) = \sigma_{\nu\mu}(x) + g_\nu(x) g_\mu(x)).$$

Proof. We apply theorem 6. Let $I(X|U)$ be equal 0 if $X < U$ and equal 1 if $X \geq U$, then

$$\Delta(i)(F_n - F) = \frac{1}{n} \sum_{s=1}^n \Delta(i)(I(\cdot | U_s) - F).$$

$U_1, \ldots, U_n$ are the sample variables, $x_j \in R^q$. Then the r. v. of theorem 6 are:

$$X_{jns} = n^{-1}(nh^q)^{1/2}h^{-q} \Sigma_i \widehat{k}^{(r)}(x_j, i)\Delta(i)\,(I(\cdot \mid U_s) - F),$$

resp.

$$n^{-1}(nh^q)^{1/2}h^{-q} \Sigma_i \widehat{k}^{(r)}(x_j, i)\,(Y_{is} - g(\xi_{is}))\Delta(i)I(\cdot \mid U_s)$$

corresponding to $f_n^{(r)}$ resp. $\gamma_n^{(r)}$.

$\mathcal{E}X_{jns} = 0.$

$\widetilde{\mathcal{C}}$ is already given in the above theorems. Only the zero's outside the block diagonal have to be proven: let be $x \neq v$

$$\mathrm{Cov}\,((nh^q)^{1/2}h^{|r|}\gamma_{n\nu}^{(r)}(x),\ (nh^q)^{1/2}h^{|r|}\gamma_{n\mu}^{(r)}(v))$$

$$= \mathcal{E}[nh^{-q}\sum_i \sum_j \widehat{k}^{(r)}(x, i) \times \widehat{k}^{(r)}(v, j) \times (\overline{y}_{(i)\nu}\Delta(i)F_n - g_\nu(\xi_i)\Delta(i)F) \times (\overline{y}_{(j)\mu}\Delta(j)F_n$$
$$- g_\mu(\xi_i)\Delta(j)F)].$$

Like in the proof of theorem 3 it follows:

$$= h^{-q} \sum_i \widehat{k}^{(r)}(x, i)\widehat{k}^{(r)}(v, j)\,[\sigma_{\nu\mu}(\eta_i)\,\Delta(i)F + g_\nu(\xi_i)g_\mu(\xi_i)\Delta(i)F(1 - \Delta(i)F)].$$

Lemma 5 implies that this sum converges to zero.

$$\mathrm{Cov}\,(nh^q)^{1/2}h^{|r|}\gamma_{n\nu}^{(r)}(x),\ (nh^q)^{1/2}h^{|r|}f_n^{(r)}(v))$$

$$= \mathcal{E}[nh^{-q}\sum_i \sum_j \widehat{k}^{(r)}(x, i) \times \widehat{k}^{(r)}(v, j) \times (\overline{y}_{(i)\nu}\,\Delta(i)\,F_n - g_\nu(\xi_i)\Delta(i)F)$$
$$\times (\Delta(j)F_n - \Delta(j)F)].$$

The proof of theorem 4 and lemma 5 show that this sum converges to zero. This proves the block diagonal form of $\mathcal{C}^*$. $\mathcal{C}^*$ is positive definite because $\widetilde{\mathcal{C}}$ is it. This is the proof of supposition (4.4.1).

Now we show (4.4.2):

$$\mathcal{E}_1(i, v, u \mid U_s) = \mathcal{E}\,[|\Delta(i)I(\cdot \mid U_s) - F| \times |\Delta(v)I(\cdot \mid U_s) - F| \times |\Delta(u)\,I(\cdot \mid U_s) - F|],$$

$$\mathcal{E}_2(i, v, u \mid U_s)$$
$$= \mathcal{E}\,[|y_{is} - g(\xi_{is})| \times |y_{vs} - g(\xi_{vs})| \times |y_{us} - g(\xi_{us})|\,\Delta(i)I(\cdot \mid U_s)\Delta(v)I(\cdot \mid U_s)\Delta(u)I(\cdot \mid U_s)].$$

There are three cases to be distinguished.

Case 1. $i = v = u$, let $p = \Delta(i)F$

$$\mathcal{E}_1\,|\Delta(i)\,(I(\cdot\, U_s) - F)\,|^3 = p(1 + O(p)), \quad \mathcal{E}_2\,|\,y_{is} - g(\xi_{is})\,|^3(\Delta(i)I(\cdot \mid U_s))^3 = \mu_3(\eta_{is}) \times p.$$

Case 2. $i = v, u \neq i$, let $p_1 = \Delta(i)F,\ p_2 = \Delta(u)F$

$$\mathcal{E}_1(i, i, u \mid U_s) = p_1 p_2(1 + O(p_1) + O(p_2)), \quad \mathcal{E}_2(i, i, u \mid U_s) = \sigma^2(\eta_{is})\mu_1(\widetilde{\eta_{is}}) \cdot 0 = 0.$$

Case 3. i, u, v different, let $p_1 = \Delta(i)F,\ p_2 = \Delta(v)F,\ p_3 = \Delta(u)F$

$$\mathcal{E}_1(i, u, v \mid U_s) = p_1 p_2 p_3(1 + O(p_1) + O(p_2) + O(p_3)), \quad \mathcal{E}_2(i, u, v \mid U_s) = 0.$$

If X_{jns} denotes a component corresponding to the density estimate:

$$\mathcal{E}|X_{jns}|^3 = (nh^q)^{-3/2}\Big[\sum_i |\widehat{k}^{(r)}(x_j, i)|^3 \Delta(i)F(1+O(\beta^q))$$

$$+3\times\sum_i\sum_k(\widehat{k}^{(r)}(x_j, i))^2|(\widehat{k}^{(r)}(x_j, u)|\Delta(i)F\Delta(u)F(1+O(\beta^q))$$

$$+\sum_i\sum_u\sum_v(\widehat{k}^{(r)}(x_j, i)\widehat{k}^{(r)}(x_j, u)k^{(r)}(x_j, u)\Delta(i)F\Delta(u)F\Delta(v)F(1+O(\beta^q)),$$

$$\mathcal{E}|X_{ins}|^3 \leq n^{-1}(nh^q)^{-1/2}[\text{const} + O(h^q)].$$

Similarly we have in the case of a component corresponding to $\gamma_n^{(r)}$

$$\mathcal{E}|X_{jns}|^3 \leq (nh^q)^{-3/2}\sum_i |\widehat{k}^{(r)}(x_j, i)|^3 \mu_3(\eta_{is})\Delta(i)F \leq n^{-1}(nh^q)^{-1/2}\,.\,\text{const}.$$

Now it follows $\sum_{s=1}^n \max_i \mathcal{E}|X_{ink}|^3 \leq (nh^q)^{-1/2}[\text{const.} + O(h^q)]$. The righthand side of this inequality converges to zero because $\lim_{n\to\infty} nh^q h^{2|r|} = \infty$ implies $\lim_{n\to\infty} nh^q = \infty$.

Theorem 8. *Let* $x \in R^q$ *with* $f^{(r)}(x) \neq 0, f(x) > 0$. *If* $\lim_{n\to\infty} nh^{q+|r|} \times \max\{b[\gamma_n^{(r)}], b[f_n^{(r)}]\} = 0$ *then* $\tilde{g}_n(x) = (nh^q)^{1/2}h^{|r|}(\gamma_n^{(r)}/f_n^{(r)} - (gf)^{(r)}/f^{(r)}$ *is asymptotically normal with expectation* 0 *and covariance matrix*

$$(\sigma_{\nu\mu}(x) + \frac{g_\nu(x)}{2}(g_\mu(x) - \frac{(gf)_\mu^{(r)}(x)}{f^{(r)}(x)}) + \frac{g_\mu(x)}{2}(g_\nu(x) - \frac{(gf)_\nu^{(r)}(x)}{f^{(r)}(x)})) \tag{4.7}$$

$$\times \frac{f(x)}{(f^{(r)}(x))^2} \times \prod_{j=1}^q \int k^{(r_j)}(y))^2 dy.$$

Proof. We consider the mapping

$$v: \begin{pmatrix} \gamma_{n1}^{(r)} \\ \vdots \\ \gamma_{nq}^{(r)} \\ f_n^{(r)} \end{pmatrix} \longrightarrow \begin{pmatrix} \gamma_{n1}^{(r)}/f_n^{(r)} \\ \vdots \\ \gamma_{nq}^{(r)}/f_n^{(r)} \end{pmatrix}.$$

According to theorem 7 $(nh^q)^{1/2}h^{|r|}Z_n^* = (nh^q)^{1/2}h^{|r|}(\gamma_{n1}^{(r)}, \ldots, \gamma_{nq}^{(r)}, f_n^{(r)})$ is asymptotic normal and Z_n^* converges in probability to $((gf)_1^{(r)}, \ldots, (gf)_q^{(r)}, f^{(r)})$ under the assumptions on h and β.

As one finds, e. g. in Witting [18, p. 49] $(nh^q)^{1/2}h^{|r|}(V(Z_n^*) - (gf)^{(r)}/f^{(r)})$ is also asymptotic normal with expectation zero and covariance matrix $\mathcal{F}((gf)^{(r)}/f^{(r)})\mathcal{C}(x)\mathcal{F}((gf)^{(r)}/f^{(r)})^T$. $\mathcal{F}((gf)^{(r)}/f^{(r)})$ is the functional determinant of the mapping V that is $\mathcal{F}(c) = (\partial V_i(c)/\partial z_j^*)$. Some computations yield (4.7).

Theorem 9. *Let* $\tilde{Z}_n = \sqrt{n}(\tilde{F}_n(x_1) - \mathcal{E}\tilde{F}_n(x_1), \ldots, \tilde{F}_n(x_1) - \mathcal{E}\tilde{F}_n(x_1))$. *It is assumed, that* $x_1, \ldots, x_l$ *are continuity points of* F *and* $F(x_j) > 0$, $j = 1, 2, \ldots, l$. *Then* $\tilde{Z}_n$ *is asymptotic normal with expectation* 0.

Proof. To prove the asymptotic normallity we apply theorem 6. $X_{jns} = n^{-1/2}\sum_i \widehat{k}^{(-1)}(x_j, i)\Delta(i)(I(\cdot|U_s) - F)$. The third moment is computed according to the proof of theorem 7 and thereof if results:

$$\mathcal{E}|X_{jns}|^3 \leqq n^{-1} n^{-1/2} \left[\sum_i |\widehat{k}^{(-1)}(x_j, i)|^3 \Delta(i) F + O(h^q)\right].$$

According to the proof of (3.2.2.) we obtain $\Sigma_i |\widehat{k}^{(-1)}(x_j, i)|^3 \Delta(i) F = \Sigma_i F(i\beta)((\widehat{k}^{(-1)}(\eta_j, i))^3)^{(\bar{1})}(\beta/\eta)^q$. This quantity converges and we get (4.4.2).

5. Minimizing the mean square error. The aim of this part is to obtain the parameter h as a function of the sample size n in order to minimize

$$\mathcal{E}(Z_n(x) - Z(x))^2 = \operatorname{Var}(Z_n(x)) + (b[Z_n(x)])^2.$$

For Z_n we take each one of the estimators $\widetilde{F}_n, f_n^{(r)}, \gamma_n^{(r)}, \gamma_n^{(r)}/f_n^{(r)}$. The above formulas of the bias show that there is not such a simple expression as in the case of the usual kernel estimate. The reason is that we use certain Riemann sums which give the additional terms of the bias.

Further we assume $\beta = h^\alpha$ with $\alpha > 0$ sufficiently large. Therefore we use only the large terms of the bias and we assume

$$(b[z_n(x)])^2 \approx B \times (h^\delta + O(h^{\delta'})) \quad \delta, \delta' > 0, \quad \delta' > \delta. \tag{5.1}$$

The variance is assumed to be

$$\operatorname{Var}(Z_n(x)) \approx A n^{-1} h^{-w}, \quad w > 0. \tag{5.2}$$

Using lemma 4a of P a r z e n [11] we get

$$\min_{h>0}(A n^{-1} h^{-w} + B h^\delta) = (\delta + w)(B/w)^w (A n^{-1}/\delta)^{\delta/(\delta+w)} \tag{5.3}$$

for $h = n^{-1/(w+\delta)}(wA/\delta B)^{1/(w+\delta)}$. If we require $\lim_{n\to\infty}(nh^q)^{1/2} h^{|r|} = \infty$ to achieve pointwise consistency then we have the condition $\varepsilon < (q + 2|r|)^{-1}$ for $h \approx C n^{-\varepsilon}$. This is fulfilled for the h in (5.3) because in this case $w = q + 2|r|$.

If we require $\lim_{n\to\infty}(nh^q)^{1/2} h^{|r|} h^{\delta/2} = 0$ to achieve asymptotic normality with expectation $\lim_{n\to\infty} \mathcal{E} Z_n(x)$ rather than $\mathcal{E} Z_n(x)$ then we have the condition $\varepsilon > (q + 2|r| + \delta)^{-1} = \varepsilon_0$ for $h \approx C n^{-\varepsilon}$. In this case we have to consider the following minimum problem

$$\min_{\varepsilon > \varepsilon_0}\left(\frac{A}{C^{wn}} n^{-1+\varepsilon w} + B C^\delta n^{-\delta\varepsilon}\right) = \min_{\varepsilon > \varepsilon_0} \widetilde{b}(\varepsilon).$$

$\widetilde{b}(\varepsilon)$ is a monoton increasing function for $\varepsilon \geqq \varepsilon_0$ and large n. Clearly the minimum is attained for $\varepsilon = \varepsilon_0$. Moreover it is possible to fulfill the two conditions $\varepsilon < (q + 2|r|)^{-1}$ and $\varepsilon > (q + 2|r| + \delta)^{-1}$ simultaneously.

6. Estimation and test of certain curve characteristics 6.1. L i n e a r i t y o f r e g r e s s i o n. We assume that the density f of X is known and that both random variables X and Y are one-dimensional. The density of X is known for example if the experimentator can choose X arbitrary and he may choose X at random in an interval $[A, B]$. Then $f(x) = (B - A)^{-1}$. In this case $\gamma_n^{(r)}/f$ is an estimator for $g^{(r)}$. If g is linear, then the second derivative is zero and $m(x) = \partial^2 \gamma_n(x)/\partial x^2$ is asymptotically normal with expectation 0. The random varfables $m(x_1), \ldots m(x_l)$ are stochastically independent in the limit. About half of the values of $m(x_1), \ldots, m(x_l)$ are greater resp. less than zero if one chooses $x_1, \ldots, x_l$ at random. The limiting independence yields now that the relative frequency of the positive values are nearly normally distributed for large l. If l tends to infinity this relative frequency is about the area of $\{x: m(x)>0, x\in[A, B]$ divided by the interval length $B-A$.

This fact may be used to test linearity. In the case of higher dimensions there have to be considered matrices of second order derivatives. Moreover, the second order derivatives may be used to research convexity of regression.

6.2 Characteristics defined by equations. For instance the mode may be defined by $\operatorname{grad} f'(\theta)=0$ and, therefore, the sample mode by $\operatorname{grad} f'_n(\theta_n)=0$. The sample fractiles are defined by $F_n(x_{\alpha,n})=\alpha$.

Let φ_n be the estimator of a certain curve φ, resp. $\varphi_n^{(r)}$ the estimator of the derivative $\varphi^{(r)}$. We consider the equations $\varphi_n^{(r)}(\theta_n)=a$ and $\varphi^{(r)}(\theta)=a$. Further we assume θ to be unique at least in a given finite interval $[A, B]$. $\varphi^{(r)}$ is assumed to be uniformly continuous.

Then the convergence of $\theta_n\to\theta$ is equivalent with the convergence of $\varphi^{(r)}(\theta_n)$ to $\varphi^{(r)}(\theta)$. The relations $|\varphi^{(r)}(\theta_n)-\varphi^{(r)}(\theta)|=|\varphi^{(r)}(\theta_n)-\varphi_n^{(r)}(\theta_n)|\leq\sup_x|\varphi^{(r)}(x)-\varphi_n^{(r)}(x)|$ show, that this convergence can be proved if $\sup_x|\varphi^{(r)}(x)-\varphi_n^{(r)}(x)|$ converges to zero (see theorem 10 below).

The asymptotic normality follows from the equation

$$a=\varphi_n^{(r_j)}(\theta_n)=\varphi_n^{(r_j)}(\theta)-\left(\frac{\partial\varphi_n^{(r_j)}}{\partial x_1}\cdots\frac{\partial\varphi_n^{(r_j)}}{\partial x_q}\right)\Bigg|_{x=\theta^*}(\theta-\theta_n),\ j=1,\ldots,q,$$

for a certain θ^*. If there exists an inverse of the matrix $(\partial\varphi_n^{(r_j)}/\partial x_i)$ then

$$\theta-\theta_n=\left(\frac{\partial\varphi_n^{(r_j)}}{\partial x_i}\right)^{-1}\Bigg|_{x=\theta^*}(\varphi_n^{(r_1)}(\theta)-a,\ldots,\varphi_n^{(r_q)}(\theta)-a)^T.$$

The asymptotic normality of the random variables $(\varphi_n^{(r_1)}(\theta)-a,\ldots,\varphi_n^{(r_q)}(\theta)-a)$ (see the above theorems) yields the asymptotic normality of $\theta-\theta_n$ if the uniform convergence in probability of $\partial\varphi_n^{(r_j)}/\partial x_i$ is proven (see theorem 10).

Theorem 10.

$$\lim_{n\to\infty}P(\sup_x|\widetilde{F}_n(x)-F(x)|>\varepsilon)=0\quad\forall\,\varepsilon>0. \tag{6.1}$$

If $\lim_{n\to\infty} nh^{2(q+|r|)}=\infty$ *then*

$$\lim_{n\to\infty}P(\sup_x|f_n^{(r)}(x)-f^{(r)}(x)|>\varepsilon)=0\quad\forall\,\varepsilon>0. \tag{6.2}$$

Proof. (6.1) and (6.2) are proved according to the same scheme:

$$\sup_x|\varphi_n(x)-\varphi(x)|\leq V_n^1+V_n^2,\ V_n^1=\sup_x|\varphi_n(x)-\mathcal{E}\varphi_n(x)|,\ V_n^2=\sup_x|\mathcal{E}\varphi_n(x)-\varphi(x)|.$$

Taking into account the proofs of the above sections V_n^2 tends to zero as $n\to\infty$ (V_n^2 is not a random variable). To prove the convergence of V_n^1 we apply lemma 1.

The structure of $\varphi_n(x)-\mathcal{E}\varphi_n(x)$ is like $\sum_i\widehat{v}(i\beta)\,\Delta(i)(F_n-F)$. By lemma 1 we obtain

$$|(-1)^q \sum_i (F_n(i\beta)-F(i\beta))\Delta(i+1)\widehat{v}| \leq (\sup_x |F_n(x)-F(x)|) \sum_i |\Delta(i+1)\widehat{v}|.$$

From Kiefer and Wolfowitz [5] we know that for the empirical distribution function of vector chance variables

$$P(\sup_x |F_n(x)-F(x)|>\lambda/n^{1/2}) \leq C_1 e^{-\alpha\lambda^2}.$$

If $\delta_n = \sup_x \Sigma_i |\Delta(i+1)\widehat{v}|$ then

$$P(V_n^1>\varepsilon) \leq P(\sup |F_n(x)-F(x)|\delta_n>\varepsilon) \leq C_1 \exp[-\alpha n\varepsilon^2/\delta_n^2].$$

Case 1. $\varphi_n = \widetilde{F}_n$ then $\delta_n \to 1$;

Case 2. $\varphi_n = f_n^{(r)}$ then $\delta_n \to h^{-q-|r|} \prod_{j=1}^q \int |k^{(r_j+1)}(y)|\, dy$.

Here of (6.1) and (6.2) are obtained.

REFERENCES

1. P. K. Bhattacharya. Estimation of a probability density function and its derivatives. *Sankhya. Indian J. Statist. Ser.* A, **29**, 1967, 373—382.
2. S. Bochner. Harmonic Analysis and the theory of probabiliзy. Univ. of California Press, 1960.
3. L. I. Boneva, D. Kendall, I. Stefanov. Spline transformations: three new diagnostic aids for the statistical data-analyst. *J. Roy. Statist. Soc., Ser.* B, **33**, 1971, 1—70.
4. T. Cacoullos. Estimation of a multivariate density. *Ann. Inst. Statist. Math.*, 18, 1966, 179—189.
5. J. Kiefer, J. Wolfowitz. On the deviations of the empiric distribution function of vector chance variables. *Trans. Amer. Math. Soc.*, **87**, 1958, 173—186.
6. M. Loève. Probability theory I. Berlin, 1977.
7. T. G. Meyer. Bounds for estimation of density functions and their derivatives. *Ann. Statist.*, 5, 1977, 136—142.
8. V. K. Murthy. Nonparametric estimation of multivariate densites with applications. In: *Multivariate analysis* (ed. P. R. Krishnaiah). New York, 1965, 43—56.
9. E. A. Nadaraya. On estimating regression. *Theory Probab. and Appl.*, 9, 1964, 141—142
10. E. A. Nadaraya. On non-parametric estimates of density functions and regression curves. *Theory Probab. and Appl.*, **10**, 1965, 186—190.
11. E. Parzen. On estimation of probability density function and mode. *Ann. Math. Statist.*, **33**, 1962, 1065—1076.
12. B. Penkov. The linear empirical spline and some applications. Symposium of applied mathematics. Thessaloniki, 1976.
13. E. F. Schuster. Joint asymptotic distribution of the estimated regression function at a finite number of distinct points. *Ann. Math. Statist.*, 43, 1972, 84—88.
14. J. Ch. Stone. Consistent nonparametric regression. *Ann. Statist.*, **5**. 1977, 595—645.
15. J. W. Tukey. Curves as parameters, and touch estimation. Proc. 4th Berkeley Symp. Math. Statistics Berkeley, 1961, 681—694.
16. G. Wahba. Optimal convergence properties of variable knot, kernel and orthogonal series methods for density estimation. *Ann. Statist.*, **3**, 1975, 15—29.
17. G. Wahba. Histosplines with knots which are order statistics. *J. Roy. Statist. Soc., Ser.* B, 38, 1976, 140—151.
18. H. Witting, G. Nölle. Angewandte Mathematische Statistik. Stuttgart, 1970.

Technische Universität
Arcisstrasse 21, D—8000 München 2

Received 31. 7. 1978