

ON SEQUENTIAL ESTIMATION OF THE REGRESSION FUNCTION

Samanta, M
Department of Statistics, The University of Manitoba

<https://doi.org/10.5109/13352>

出版情報 : Bulletin of informatics and cybernetics. 21 (1/2), pp.19-27, 1984-03. Research
Association of Statistical Sciences
バージョン :
権利関係 :

ON SEQUENTIAL ESTIMATION OF THE REGRESSION FUNCTION*

By

M. SAMANTA**

1. Introduction and Summary

Let (X, Y) be a two dimensional random variable having a joint density function f and let g be the marginal density function of X . We assume that $E_f Y$ is finite and define the regression function $m(x)$ (for regression of Y on X) by $m(x) = E[Y|X=x]$. Nadaraya [3] and Watson [7] and Schuster [6] have studied the asymptotic properties of the estimate $\hat{m}_n(x)$ of $m(x)$ defined by

$$(1) \quad \hat{m}_n(x) = \frac{\sum_{i=1}^n Y_i K\left(\frac{x-X_i}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)}$$

where $K(u)$ is a probability density function on $(-\infty, +\infty)$, $\{h_n\}$ is a monotonically decreasing sequence of positive numbers converging to zero and $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ are n independent observations of (X, Y) .

In many practical situations the number of observations N_t which we observe in time $(0, t]$ is a random variable. We call N_t a stopping random variable. We assume that $(X_1, Y_1), (X_2, Y_2), \dots$ are independent observations of (X, Y) and need not be independent of the random variable N_t . In this paper we propose an estimate $m_{N_t}(x)$ of the regression function $m(x)$ based on $(X_1, Y_1), (X_2, Y_2), \dots, (X_{N_t}, Y_{N_t})$ and given by

$$(2) \quad m_{N_t}(x) = \frac{\sum_{i=1}^{N_t} Y_i K\left(\frac{x-X_i}{h_i}\right)}{\sum_{i=1}^{N_t} \frac{1}{h_i} K\left(\frac{x-X_i}{h_i}\right)}$$

We note that the expression for the estimate $m_{N_t}(x)$ is motivated by the recursive type of estimate of a probability density function first proposed by Yamato [8]. Suppose x_1, x_2, \dots, x_l are l distinct points. We have shown that under certain regularity conditions $(N_t h_{N_t})^{1/2} \{m_{N_t}(x_1) - m(x_1), \dots, m_{N_t}(x_l) - m(x_l)\}$ is asymptotically normally distributed with mean vector 0 and diagonal covariance matrix $C = [c_{ij}]$ with $c_{ii} = \frac{\text{Var}[Y|X=x_i]}{g(x_i)} \nu \int_{-\infty}^{\infty} K^2(u) du$ where $\nu (\nu < 1)$ is as defined in the next section. For simplicity we have proved the theorem for the special case $l=2$. The method of proof remains valid in the more general case. The theorem can be regarded as the appropriate

* Research partially supported by a Natural Science and Enging Research Council Grant of Canada.

** Department of Statistics, The University of Manitoba, Winnipeg, Manitoba, R3t 2n2 Canada.

extension of the earlier result due to Schuster [6].

2. Main Result

We assume that the probability density function K and the sequence $\{h_n\}$ are chosen to satisfy the following conditions:

- (i) $K(u)$ and $|uK(u)|$ are bounded
- (ii) $\int_{-\infty}^{\infty} uK(u)du=0$
- (iii) $\int_{-\infty}^{\infty} u^2K(u)du<\infty$
- (iv) $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \left(\frac{h_n}{h_j} \right) = \nu < 1$
- (v) $\frac{1}{n} \sum_{j=1}^n h_j^2 \leq C_1 h_n^2$ ($C_1 > 0$), $n=1, 2, 3, \dots$
- (vi) $\lim_{n \rightarrow \infty} nh_n^3 = \infty$ and $\lim_{n \rightarrow \infty} nh_n^5 = 0$

REMARK 1. If K is the standard normal probability density function, then conditions (i), (ii) and (iii) are satisfied.

REMARK 2. If $h_n = n^{-\delta}$, $\frac{1}{5} < \delta < \frac{1}{3}$, then conditions (iv), (v) and (vi) are satisfied. (For a proof see [1], p. 26 and p. 46.)

We define the following:

$$\begin{aligned} g(x) &= \int_{-\infty}^{\infty} f(x, y) dy \\ (3) \quad w(x) &= \int_{-\infty}^{\infty} y f(x, y) dy \\ v(x) &= \int_{-\infty}^{\infty} y^2 f(x, y) dy. \end{aligned}$$

Hence,

$$\text{Var}[Y|X=x] = \frac{v(x)}{g(x)} - \frac{w^2(x)}{g^2(x)}.$$

Suppose x_1 and x_2 are two distinct points. We define for $i=1, 2, \dots, n$ and $s=1, 2$

$$\begin{aligned} U_i^*(x_s) &= \frac{1}{h_i} K\left(\frac{x_s - X_i}{h_i}\right) \\ V_i^*(x_s) &= Y_i U_i^*(x_s) \\ U_i(x_s) &= h_i^{1/2} \{U_i^*(x_s) - EU_i^*(x_s)\} \\ V_i(x_s) &= h_i^{1/2} \{V_i^*(x_s) - EV_i^*(x_s)\} \\ \bar{U}_n(x_s) &= \sum_{i=1}^n U_i(x_s) \end{aligned}$$

$$\begin{aligned}
(4) \quad \bar{V}_n(x_s) &= \sum_{i=1}^n V_i(x_s) \\
W_i &= (U_i(x_1), V_i(x_1), U_i(x_2), V_i(x_2))' \\
n^{1/2} Z_n &= (\bar{U}_n(x_1), \bar{V}_n(x_1), \bar{U}_n(x_2), \bar{V}_n(x_2))' \\
A &= \nu \int_{-\infty}^{\infty} K^2(u) du \begin{pmatrix} g(x_1) & w(x_1) & 0 & 0 \\ w(x_1) & v(x_1) & 0 & 0 \\ 0 & 0 & g(x_2) & w(x_2) \\ 0 & 0 & w(x_2) & v(x_2) \end{pmatrix}
\end{aligned}$$

Let Z be a four variate normal random variable with mean vector 0 and covariance matrix A .

We now prove the following lemmas.

LEMMA 1. Suppose K satisfies conditions (i) and (iii) and the sequence $\{h_n\}$ satisfies condition (iv). Let g' , w' and v' exist and be bounded. Then the following results hold for $s=1, 2$ and $r=1, 2$.

- (a) $\lim_{n \rightarrow \infty} \text{Var} \{n^{-1/2} \bar{U}_n(x_s)\} = \nu g(x_s) \int_{-\infty}^{\infty} K^2(u) du$
- (b) $\lim_{n \rightarrow \infty} \text{Var} \{n^{-1/2} \bar{V}_n(x_s)\} = \nu v(x_s) \int_{-\infty}^{\infty} K^2(u) du$
- (c) $\lim_{n \rightarrow \infty} \text{Cov} \{n^{-1/2} \bar{U}_n(x_s), n^{-1/2} \bar{V}_n(x_s)\} = \nu w(x_s) \int_{-\infty}^{\infty} K^2(u) du$
- (d) $\lim_{n \rightarrow \infty} \text{Cov} \{n^{-1/2} \bar{U}_n(x_1), n^{-1/2} \bar{U}_n(x_2)\} = 0$
- (e) $\lim_{n \rightarrow \infty} \text{Cov} \{n^{-1/2} \bar{V}_n(x_1), n^{-1/2} \bar{V}_n(x_2)\} = 0$
- (f) $\lim_{n \rightarrow \infty} \text{Cov} \{n^{-1/2} \bar{U}_n(x_s), n^{-1/2} \bar{V}_n(x_r)\} = 0, \quad r \neq s.$

PROOF. We sketch the proof of part (a) and part (d) of the Lemma. The proof of the other parts are similar and will be omitted. To obtain part (a) we have

$$\begin{aligned}
\text{Var} \{n^{-1/2} \bar{U}_n(x_s)\} &= \frac{1}{n} \sum_{j=1}^n \left(\frac{h_n}{h_j} \right) \int_{-\infty}^{\infty} \{K(u)\}^2 g(x_s - h_j u) du \\
&\quad - \frac{h_n}{n} \sum_{j=1}^n \left\{ \int_{-\infty}^{\infty} K(u) g(x_s - h_j u) du \right\}^2 \\
&= \frac{1}{n} \sum_{j=1}^n \left(\frac{h_n}{h_j} \right) \left[g(x_s) \int_{-\infty}^{\infty} \{K(u)\}^2 du + O(h_j) \right] \\
&\quad - \frac{h_n}{n} \sum_{j=1}^n [g(x_s) + O(h)]_j.
\end{aligned}$$

Hence,

$$\lim_{n \rightarrow \infty} \text{Var} \{n^{-1/2} \bar{U}_n(x_s)\} = \nu g(x_s) \int_{-\infty}^{\infty} K^2(u) du, \quad s=1, 2.$$

To prove part (d) we have used the method similar to that of Schuster [6]:

$$E\left[\frac{1}{h_i}K\left(\frac{x_1-X_i}{h_i}\right)K\left(\frac{x_2-X_i}{h_i}\right)\right]=O(h_i).$$

Now,

$$\begin{aligned}\text{Cov}\{n^{-1/2}\bar{U}_n(x_1), n^{-1/2}\bar{U}_n(x_2)\} &= \frac{1}{n} \sum_{i=1}^n E\{U_i(x_1)U_i(x_2)\} \\ &= \frac{1}{n} \sum_{i=1}^n \left[\frac{h_n}{h_i} E\left\{\frac{1}{h_i}K\left(\frac{x_1-X_i}{h_i}\right)K\left(\frac{x_2-X_i}{h_i}\right)\right\} \right. \\ &\quad \left. - h_n E\left\{\frac{1}{h_i}K\left(\frac{x_1-X_i}{h_i}\right)\right\} E\left\{\frac{1}{h_i}K\left(\frac{x_2-X_i}{h_i}\right)\right\} \right] \\ &= \frac{1}{n} \sum_{i=1}^n \left[\frac{h_n}{h_i} \cdot O(h_i) - h_n \{g(x_1)+O(h_i)\} \{g(x_2)+O(h_i)\} \right] \\ &= O(h_n).\end{aligned}$$

Hence,

$$\lim_{n \rightarrow \infty} \text{Cov}\{n^{-1/2}\bar{U}_n(x_1), n^{-1/2}\bar{U}_n(x_2)\} = 0.$$

Let $C=(c_1, d_1, c_2, d_2)'$ be any real vector in R^4 .

LEMMA 2. Suppose K satisfies conditions (i) and (iii) and the sequence $\{h_n\}$ satisfies condition (iv) and $nh_n^3 \rightarrow \infty$ as $n \rightarrow \infty$. Let $E_f|Y|^3$ be finite and let g', w' and v' exist and be bounded. If $g(x_i) > 0$ for $i=1, 2$, then $C'Z_n$ converges in distribution to a normal random variable with mean 0 and variance $C'AC$.

PROOF. We shall establish the asymptotic normality of $C'Z_n$ by showing that

$$(5) \quad \lim_{n \rightarrow \infty} \left\{ \frac{\sum_{i=1}^n E|C'W_i|^3}{n^{3/2}(\text{Var}(C'Z_n))^{3/2}} \right\} = 0. \quad (\text{See [2], p. 275.})$$

Using Lemma 1, we have

$$(6) \quad \begin{aligned}\lim_{n \rightarrow \infty} \text{Var}(C'Z_n) &= \nu \int_{-\infty}^{\infty} K^2(u) du \left[\sum_{s=1}^2 \{c_s^2 g(x_s) + d_s^2 v(x_s) + 2c_s d_s w(x_s)\} \right] \\ &= C'AC > 0.\end{aligned}$$

The positive definiteness of the matrix A follows from the fact that $g(x)v(x) - w^2(x) = g^2(x)\text{Var}[Y|X=x]$ and $g(x_i) > 0$ for $i=1, 2$. It now suffices to prove that

$$\lim_{n \rightarrow \infty} n^{-3/2} \left\{ \sum_{i=1}^n E|C'W_i|^3 \right\} = 0$$

Using the hypothesis and the arguments similar to those in Lemma 1 it can be shown that

$$E|U_i(x_s)|^3 = O(h_n^{3/2} h_i^{-2}) = O(h_n^{-1/2})$$

and

$$E|V_i(x_s)|^3 = O(h_n^{3/2} h_i^{-3}) = O(h_n^{-3/2}).$$

Hence

$$\begin{aligned}
n^{-3/2} \left\{ \sum_{i=1}^n E |C'W_i|^3 \right\} &\leq n^{-3/2} |C|^3 \sum_{i=1}^n E |W_i|^3 \\
&\leq 8n^{-3/2} |C|^3 \sum_{i=1}^n \max_{s=1,2} \{E |U_i(x_s)|^3, E |V_i(x_s)|^3\} \\
&= O \{ (nh_n^3)^{-1/2} \} = o(1).
\end{aligned}$$

This completes the proof.

We define

$$\begin{aligned}
(7) \quad Z_n^* &= h_n^{1/2} \cdot n^{-1/2} \left\{ \sum_{i=1}^n (U_i^*(x_1) - g(x_1)), \sum_{i=1}^n (V_i^*(x_1) - w(x_1)), \right. \\
&\quad \left. \sum_{i=1}^n (U_i^*(x_2) - g(x_2)), \sum_{i=1}^n (V_i^*(x_2) - w(x_2)) \right\}.
\end{aligned}$$

LEMMA 3. Suppose conditions (i) through (vi) are satisfied. Let $E_f |Y|^3$ be finite and let g', g'', w', w'' and v' exist and be bounded. If $g(x_i) > 0$ for $i=1, 2$, then $C'Z_n^*$ converges in distribution to a normal random variable with mean 0 and variance $C'AC$.

PROOF. We have

$$(8) \quad C'Z_n - C'Z_n^* = h_n^{1/2} \cdot n^{-1/2} \sum_{i=1}^n \sum_{s=1}^2 [c_s \{g(x_s) - EU_i^*(x_s)\} + d_s \{w(x_s) - EV_i^*(x_s)\}].$$

Using the hypothesis it can be shown that for $i=1, 2, \dots, n$ and $s=1, 2$

$$EU_i^*(x_s) = g(x_s) + O(h_i^2)$$

and

$$EV_i^*(x_s) = w(x_s) + O(h_i^2).$$

Hence,

$$(9) \quad \left| \sum_{i=1}^n (g(x_s) - EU_i^*(x_s)) \right| = O \left(\sum_{i=1}^n h_i^2 \right),$$

and

$$(10) \quad \left| \sum_{i=1}^n (w(x_s) - EV_i^*(x_s)) \right| = O \left(\sum_{i=1}^n h_i^2 \right).$$

From (8), (9) and (10) we get

$$\begin{aligned}
C'Z_n - C'Z_n^* &= O \left(h_n^{1/2} n^{-1/2} \sum_{i=1}^n h_i^2 \right) \\
&= O \left((nh_n)^{1/2} \frac{1}{n} \sum_{i=1}^n h_i^2 \right) \\
&= O((nh_n)^{1/2} h_n^2) \\
&= O((nh_n^5)^{1/2}) \\
&= o(1).
\end{aligned}$$

The proof now follows from Lemma 2.

Suppose N_t ($t > 0$) is a stopping random variable such that $\frac{N_t}{t} \xrightarrow{p} \pi$ ($\pi > 0$) as $t \rightarrow \infty$.

This implies that for any $\varepsilon > 0$, there exists $t_0 = t_0(\varepsilon)$ such that for $t \geq t_0$ we have

$$P[|N_t - \pi t| \geq \pi t \varepsilon] < \varepsilon.$$

We define

$$(11) \quad N_1 = N_1(t, \varepsilon) = [\pi t(1 - \varepsilon)]$$

and

$$(12) \quad N_2 = N_2(t, \varepsilon) = [\pi t(1 + \varepsilon)]$$

where $[x]$ is the integral part of x .

We note that for any $0 < \varepsilon < \frac{1}{2}$ and $t > \frac{1}{\pi \varepsilon}$ the numbers N_1 and N_2 defined above satisfy the following inequalities:

$$\frac{N_2}{N_1} < \frac{1 + \varepsilon}{1 - 2\varepsilon}$$

and

$$\frac{N_2 - N_1}{N_1} < \frac{3\varepsilon}{1 - 2\varepsilon}.$$

We define

$$R_i = \sum_{s=1}^2 \{c_s U_i^*(x_s) + d_s V_i^*(x_s)\}, \quad i=1, 2, \dots$$

$$\mu = \sum_{s=1}^2 \{c_s g(x_s) + d_s w(x_s)\}$$

$$S_n = \sum_{i=1}^n \{R_i - \mu\}, \quad n=1, 2, \dots$$

$$Q = \max_{N_1 < n \leq N_2} \left| \sum_{i=N_1+1}^n [R_i - E\{R_i\}] \right|.$$

It can be seen that $h_n^{1/2} n^{-1/2} S_n = C' Z_n^*$. Replacing n by N_t in the expressions for $\bar{U}_n(x_s)$, $\bar{V}_n(x_s)$, Z_n and Z_n^* , we define $\bar{U}_{N_t}(x_s)$, $\bar{V}_{N_t}(x_s)$, Z_{N_t} and $Z_{N_t}^*$ respectively.

In order to study the asymptotic distribution of $C' Z_{N_t}^*$ we find it convenient to choose a specific sequence $\{h_n = n^{-\delta}, n=1, 2, \dots\}$, where δ is some positive number. With this choice of $\{h_n\}$ we have the following lemma. Let C_1 be a generic constant.

LEMMA 4. Suppose K satisfies conditions (i) and (iii) and $\{h_n = n^{-\delta}\}$, $\delta > 0$. If g' and v' exist and are bounded, then for any $0 < \varepsilon < \frac{1}{2}$, $t > \frac{1}{\pi \varepsilon}$ we have

$$P\left\{Q \geq \varepsilon^{1/3} \sqrt{\frac{N_1}{h_{N_1}}}\right\} < \frac{C_1 \varepsilon^{1/3}}{(1-2\varepsilon)} \left\{ \frac{(1+\varepsilon)}{(1-2\varepsilon)} \right\}^\delta$$

PROOF. By Kolmogorov's inequality

$$(13) \quad P\left[Q \geq \varepsilon^{1/3} \sqrt{\frac{N_1}{h_{N_1}}}\right] \leq \frac{\sum_{i=N_1+1}^{N_2} E\{R_i^2\}}{\varepsilon^{2/3} \left\{ \frac{N_1}{h_{N_1}} \right\}}.$$

Using the hypothesis it can be shown that

$$(14) \quad E\{R_i^2\} = O\left(\frac{1}{h_i}\right) = O\left(\frac{1}{h_{N_2}}\right) \quad \text{if } N_1 < i \leq N_2.$$

From (13) and (14), we get

$$\begin{aligned} P\left[Q \geq \varepsilon^{1/3} \sqrt{\frac{N_1}{h_{N_1}}}\right] &\leq \frac{C_1(N_2 - N_1)}{3\varepsilon^{2/3}N_1} \left(\frac{h_{N_1}}{h_{N_2}}\right) \\ &\leq \frac{C_1\varepsilon^{1/3}}{(1-2\varepsilon)} \left(\frac{1+\varepsilon}{(1-2\varepsilon)}\right)^\delta. \end{aligned}$$

LEMMA 5. Suppose K satisfies conditions (ii), and (iii) and $\{h_n = n^{-\delta}\}$, $\delta > 0$. If g' , g'' , w' and w'' exist and are bounded and if $0 < \varepsilon < \frac{1}{2}$, $t > \frac{1}{\pi\varepsilon}$ and $\frac{1}{5} < \delta$, then for all $N_1 < n \leq N_2$,

$$\left| \sqrt{\frac{h_{N_1}}{N_1}} \sum_{i=N_1+1}^n [E\{R_i\} - \mu] \right| < \frac{C_1\varepsilon}{(1-2\varepsilon)}.$$

PROOF. Using computations similar to those in Lemma 3, we obtain for all n such that $N_1 < n \leq N_2$

$$\begin{aligned} \left| \sum_{i=N_1+1}^n [E\{R_i\} - \mu] \right| &\leq \frac{C_1}{3} (N_2 - N_1) \sum_{i=N_1+1}^n h_i^2 \\ &\leq \frac{C_1}{3} (N_2 - N_1) h_{N_1}^2. \end{aligned}$$

Hence,

$$\begin{aligned} \left| \sqrt{\frac{h_{N_1}}{N_1}} \sum_{i=N_1+1}^n [E\{R_i\} - \mu] \right| &\leq \frac{C_1(N_2 - N_1)}{3N_1} (N_1 h_{N_1}^5)^{1/2} \\ &= \frac{C_1(N_2 - N_1)}{3N_1} (N_1^{1-5\delta})^{1/2} \\ &\leq \frac{C_1\varepsilon}{(1-2\varepsilon)}. \end{aligned}$$

LEMMA 6. Suppose K satisfies conditions (i), (ii) and (iii) and $\{h_n = n^{-\delta}\}$, $\frac{1}{5} < \delta < \frac{1}{3}$ and $\frac{N_t}{t} \xrightarrow{P} \pi$ ($\pi > 0$) as $t \rightarrow \infty$. Let $E_f|Y|^3$ be finite and let g' , g'' , w' , w'' and v' exist and be bounded. If $x_1 \neq x_2$, $g(x_i) > 0$ for $i=1, 2$, then $C'Z_{N_t}^*$ converges in distribution to a normal random variable with mean 0 and variance $C'AC$ as t tends to infinity.

PROOF. The proof resembles that of Theorem 1 in Renyi [5]. Let ε ($\varepsilon < \frac{1}{2}$) be an arbitrarily small positive number. Let $t \geq t_0$ where $t_0 = t_0(\varepsilon) > \frac{1}{\pi\varepsilon}$ and let N_1 and N_2 be chosen as before.

We have for $y > 0$

$$\begin{aligned} P[C'Z_{N_t}^* < y] &= \sum_{n=1}^{\infty} P[C'Z_n^* < y; N_t = n] \\ &= \sum_{|n - \pi t| < \pi t\varepsilon} P[C'Z_n^* < y; N_t = n] \\ &\quad + \sum_{|n - \pi t| \geq \pi t\varepsilon} P[C'Z_n^* < y; N_t = n] \\ &\leq \sum_{|n - \pi t| < \pi t\varepsilon} P[C'Z_n^* < y; N_t = n] + \varepsilon. \end{aligned}$$

Hence,

$$(15) \quad |P[C'Z_{N_t}^* < y] - \sum_{|n - \pi t| < \pi t \varepsilon} P[C'Z_n^* < y; N_t = n]| < \varepsilon.$$

Introducing the random variables R_i , S_n and Q as defined in (11), we have for any n such that $N_1 < n \leq N_2$

$$(16) \quad \begin{aligned} P[C'Z_n^* < y; N_t = n] &= P\left[\sqrt{\frac{h_n}{n}} \left\{ \sum_{i=1}^{N_1} (R_i - \mu) + \sum_{i=N_1+1}^n (R_i - E(R_i)) \right. \right. \\ &\quad \left. \left. + \sum_{i=N_1+1}^n (E(R_i) - \mu) \right\} < y; N_t = n\right] \\ &\leq P\left[S_{N_1} < y \sqrt{\frac{N_2}{h_{N_2}}} - \sum_{i=N_1+1}^n (E(R_i) - \mu) + Q; N_t = n\right] \\ &= P\left[\sqrt{\frac{h_{N_1}}{N_1}} S_{N_1} < y \sqrt{\frac{N_2}{N_1} \left(\frac{h_{N_1}}{h_{N_2}}\right)} - \sqrt{\frac{h_{N_1}}{N_1}} \sum_{i=N_1+1}^n \{E(R_i) - \mu\} \right. \\ &\quad \left. + \sqrt{\frac{h_{N_1}}{N_1}} Q; N_t = n\right] \\ &\leq P\left[C'Z_{N_1}^* < y \sqrt{\left(\frac{N_2}{N_1}\right)^{1+\delta}} + \frac{C_1 \varepsilon}{(1-2\varepsilon)} + \sqrt{\frac{h_{N_1}}{N_1}} Q; N_t = n\right], \end{aligned}$$

by Lemma 5. Hence,

$$(17) \quad \begin{aligned} &\sum_{|n - \pi t| < \pi t \varepsilon} P[C'Z_n^* < y; N_t = n] \\ &\leq P\left[C'Z_{N_1}^* < y \sqrt{\left(\frac{N_2}{N_1}\right)^{1+\delta}} + \frac{C_1 \varepsilon}{(1-2\varepsilon)} + \sqrt{\frac{h_{N_1}}{N_1}} Q; |N_t - \pi t| < \pi t \varepsilon\right] \\ &\leq P\left[C'Z_{N_1}^* < y \left(\frac{1+\varepsilon}{(1-2\varepsilon)}\right)^{(1+\delta)/2} + \frac{C_1 \varepsilon}{(1-2\varepsilon)} + \varepsilon^{1/3}; \right. \\ &\quad \left. Q < \varepsilon^{1/3} \sqrt{\frac{N_1}{h_{N_1}}}; |N_t - \pi t| < \pi t \varepsilon\right] + P\left[Q \geq \varepsilon^{1/3} \sqrt{\frac{N_1}{h_{N_1}}}\right] \\ &\leq P\left[C'Z_{N_1}^* < y \left(\frac{1+\varepsilon}{(1-2\varepsilon)}\right)^{(1+\delta)/2} + \frac{C_1 \varepsilon}{(1-2\varepsilon)} + \varepsilon^{1/3}\right] + \frac{C_1 \varepsilon^{1/3}}{(1-2\varepsilon)} \left(\frac{1+\varepsilon}{(1-2\varepsilon)}\right)^\delta, \end{aligned}$$

by Lemma 4.

From (16) we get in a similar manner

$$(18) \quad \begin{aligned} &\sum_{|n - \pi t| < \pi t \varepsilon} P[C'Z_n^* < y; N_t = n] \\ &\geq P\left[C'Z_{N_1}^* < y - \frac{C_1 \varepsilon}{(1-2\varepsilon)} - \sqrt{\frac{h_{N_1}}{N_1}} Q; |N_t - \pi t| < \pi t \varepsilon\right] \\ &\geq P\left[C'Z_{N_1}^* < y - \frac{C_1 \varepsilon}{(1-2\varepsilon)} - \varepsilon^{1/3}; Q < \varepsilon^{1/3} \sqrt{\frac{N_1}{h_{N_1}}}; |N_t - \pi t| < \pi t \varepsilon\right] \\ &\geq P\left[C'Z_{N_1}^* < y - \frac{C_1 \varepsilon}{(1-2\varepsilon)} - \varepsilon^{1/3}\right] - P\left[Q \geq \varepsilon^{1/3} \sqrt{\frac{N_1}{h_{N_1}}}\right] - P[|N_t - \pi t| \geq \pi t \varepsilon] \\ &\geq P\left[C'Z_{N_1}^* < y - \frac{C_1 \varepsilon}{(1-2\varepsilon)} - \varepsilon^{1/3}\right] - \frac{C_1 \varepsilon^{1/3}}{(1-2\varepsilon)} \left(\frac{1+\varepsilon}{(1-2\varepsilon)}\right)^\delta - \varepsilon. \end{aligned}$$

From (15), (17) and (18) we conclude that for $t \geq t_0$

$$P[C'Z_{N_t}^* < y] \leq P\left[C'Z_{N_1}^* < y\left(\frac{1+\varepsilon}{1-2\varepsilon}\right)^{(1+\delta)/2} + \frac{C_1\varepsilon}{(1-2\varepsilon)} + \varepsilon^{1/3}\right] + \frac{C_1\varepsilon^{1/3}}{(1-2\varepsilon)}\left(\frac{1+\varepsilon}{1-2\varepsilon}\right)^{\delta} + \varepsilon$$

and

$$P[C'Z_{N_t}^* < y] \geq P\left[C'Z_{N_1}^* < y - \frac{C_1\varepsilon}{(1-2\varepsilon)} - \varepsilon^{1/3}\right] - \frac{C_1\varepsilon^{1/3}}{(1-2\varepsilon)}\left(\frac{1+\varepsilon}{1-2\varepsilon}\right)^{\delta} - 2\varepsilon.$$

Similar statements hold for $y < 0$. We now invoke Lemma 3 and the continuity of the distribution function of a normal random variable to complete the proof.

We are now in a position to prove the main theorem of this paper.

THEOREM. Suppose K satisfies conditions (i), (ii) and (iii), and $\{h_n = n^{-\delta}\}$, $\frac{1}{5} < \delta < \frac{1}{3}$ and $\frac{N_t}{t} \xrightarrow{p} \pi$ ($\pi > 0$) as $t \rightarrow \infty$. Let $E_f|Y|^3$ be finite and let g', g'', w', w'' and v' exist and be bounded. If $x_1 \neq x_2$, $g(x_i) > 0$ for $i=1, 2$, then $(N_t h_{N_t})^{1/2}(m_{N_t}(x_1) - m(x_1), m_{N_t}(x_2) - m(x_2))'$ converges in distribution to Z^* as t tends to infinity where Z^* is a bivariate normal random variable with mean vector 0 and diagonal covariance matrix $C = [c_{ij}]$ where

$$c_{ii} = \frac{\text{Var}[Y|X=x_i]}{g(x_i)} \nu \int_{-\infty}^{\infty} K^2(u) du \quad i=1, 2.$$

PROOF. Using the Cramér-Wold theorem (Theorem (xi) on page 123 of [4]) we conclude from Lemma 6 that $Z_{N_t}^*$ converges in distribution to Z as t tends to infinity. The proof of the theorem now immediately follows from this result in conjunction with Theorem (iii) on page 388 of [4].

References

- [1] KOROVKIN, P.P.: Inequalities. Pergamon Press, London, (1961).
- [2] LOEVE, M.: Probability Theory. Third Edition, Van Nostrand, Princeton, (1963).
- [3] NADARAYA, E.A.: "On Estimating Regression". Theory of Probability and Its Applications, **9** (1964), 141-142.
- [4] RAO, C.R.: Linear Statistical Inference and Its Applications. Second Edition, Wiley, New York, (1973).
- [5] RENYI, A.: On the Asymptotic Distribution of the Sum of a Random Number of Independent Random Variables. Acta Mathematica Scientiarum Hungaricae, **8** (1957), 193-199.
- [6] SCHUSTER, E.F.: Joint Asymptotic Distribution of the Estimated Regression at a Finite Number of Distinct Points. Annals of Mathematical Statistics, **43** (1972), 84-88.
- [7] WATSON, G.S.: "Smooth Regression Analysis". Sankhyā, **A26** (1964), 359-372.
- [8] YAMATO, H.: "Sequential Estimation of a Continuous Probability Density Function and Mode". Bulletin of Mathematical Statistics, **14**, 3-4, (1971), 1-12.

Communicated by S. Kano

Received June 20, 1983