by

Steinar Bjerve ${ }^{1}$<br>University of Oslo,<br>Kjell A. Doksum ${ }^{2}$<br>Department of Statistics University of California, Berkeley

and
Brian S. Yandell
Statistics Department
University of Wisconsin


#### Abstract

Let $(X, Y)$ be a bivariate random variable and let $m(x)=E(Y \mid X=X)$ be the regression function of $Y$ on $X$. Suppose that $Y_{1}, \ldots, Y_{n}$ are independent observations of $Y$ at $X=x_{1}, \ldots, X_{n}$. We consider nearest neighbor estimates, $\hat{m}(x)$, and employ well-known inequalities to obtain exact and asymptotic uniform confidence bounds for $E \hat{m}(x)$ and $m(x)$ based on $\hat{m}(x)$. Finally we discuss bias-properties of $\hat{m}(x)$.


Key words: Nonparametric regression, confidence bands, nearest neighbor.

[^0]1. Introduction.

Let $\left(X_{1}, Y_{1}\right), \ldots,\left(X_{n}, Y_{n}\right)$ be a random sample from a bivariate population with distribution function $F(x, y)$. We are interested in constructing uniform confidence bounds for the unknown regression function,

$$
m(x)=E(Y \mid X=x)
$$

without making parametric assumptions about either $m$ or the distributional form of $F$. We will assume existence of the conditional variance function given by

$$
\sigma^{2}(x)=\operatorname{var}(Y \mid X=x) .
$$

The construction will be based on the k-nearest neighbor estimator,

$$
\hat{m}(x)=\Sigma Y_{i} / k,
$$

where the summation is taken over the indexes of the $k X$ 's that lie closest to x .

## 2. Prelimiaries.

Let $X_{n i}$ be the ith order statistic obtained from $X_{1}, X_{2}, \ldots, X_{n}$ and let $Y_{n i}$ denote the ith induced order statistic of the $Y$-observations. I.e. if $Q_{i}$ is the anti-rank of $X_{i}, X_{Q_{i}}=X_{n i}$, then $Y_{n i}=Y_{Q_{i}}$. Conditional on $X_{n i}=x_{n i}, i=1, \ldots, n$, the $Y_{n i}$ are independent (see e.g. Bhattacharya (1974), Lemma 1) and the distributional assumptions may be written as follows:

$$
y_{n i}=m\left(x_{n i}\right)+\epsilon_{n i} \quad i=1, \ldots, n,
$$

where $\epsilon_{n 1}, \ldots, \epsilon_{n n}$ are independent,

$$
x_{n 1} \leqslant \ldots \leqslant x_{n n}, \quad E \epsilon_{n i}=0 \text { and } \operatorname{Var}\left(\epsilon_{n i}\right)=\sigma^{2}\left(x_{n i}\right) .
$$

For convenience we will, until Section 5, write $Y_{i}$ for $Y_{n i}$ and $x_{i}$ for $x_{n i}$.

Our estimator for $m(x)$ will be the $k$-nearest neighbor estimator,

$$
\begin{equation*}
\hat{m}(x)=\Sigma_{i \in I_{n k}}(x)^{\left\{Y_{i} / k\right\}}, \tag{2.2}
\end{equation*}
$$

where $I_{n k}(x)$ are the indices of the $k$ values of $x_{1}, \ldots, x_{n}$ closest to $x$. We assume that $I_{n k}(x)$ is uniquely determined a.s.

Let

$$
\begin{equation*}
J_{i}=\left\{x: I_{n k}(x)=\{i+1, \ldots, i+k\}\right\}, \quad i=0, \ldots, n-k . \tag{2.3}
\end{equation*}
$$

Denote

$$
\hat{m}_{i}=\hat{m}(x) \text { for } x \in J_{i}, \quad i=0, \ldots, n-k
$$

Thus

$$
\begin{equation*}
\hat{m}_{i}=\left(Y_{i+1}+\cdots+Y_{i+k}\right) / k, \quad i=0, \ldots, n-k . \tag{2.4}
\end{equation*}
$$

It is seen that $J_{0}=\left(-\infty, \frac{1}{2}\left(x_{1}+x_{k+1}\right)\right)$ while

$$
u_{i}=\left(\frac{1}{2}\left(x_{i}+x_{i+k}\right), \frac{1}{2}\left(x_{i+1}+x_{i+k+1}\right)\right), \quad i=1, \ldots, n-k .
$$


Let

$$
S_{i}=\hat{m}_{i}-E \hat{m}_{i}=\Sigma_{j=i+1}^{i+k} W_{j}, \quad i=0, \ldots, n-k
$$

where

$$
W_{j}=\left(Y_{j}-m\left(x_{j}\right)\right) / k, \quad j=1, \ldots, n
$$

Define $\widetilde{S}_{0}=0$ and

$$
\widetilde{S}_{i}=\Sigma_{j=1}^{i} W_{j}, \quad i=1, \ldots, n-k
$$

Then

$$
s_{i}=\tilde{s}_{i+k}-\tilde{S}_{i}, \quad i=0, \ldots, n-k
$$

LEMMA. For each $t>0$,
$P\left(S_{i} \geqslant-t ;\right.$ all $\left.i=0, \cdots, n-k\right) \geq 1-4(k t)^{-2}\left[\sum_{i=1}^{n-k} \sigma^{2}\left(x_{i}\right)+\sum_{i=k}^{n} \sigma^{2}\left(x_{i}\right)\right]$.
PROOF. Let

$$
\begin{aligned}
& A=\left\{\tilde{S}_{k+i} \geq-\frac{1}{2} t \quad \text { all } \quad i=0, \cdots, n-k\right\} \\
& B=\left\{\tilde{S}_{i} \leq \frac{1}{2} t \text { all } i=1, \cdots, n-k\right\} \\
& C=\left\{S_{i} \geq-t \text { all } i=0, \cdots, n-k\right\}
\end{aligned}
$$

Then $A \cap B \subset C$ and by Bonferroni's inequality, we have

$$
P(C) \geq P(A \cap B) \geq 1-P\left(A^{C}\right)-P\left(B^{C}\right)
$$

Kolmogorov's inequality (e.g., Loève, 1963, p. 235), yields

$$
\begin{aligned}
& P(A) \geq 1-4(k t)^{-2} \sum_{i=k}^{n} \sigma^{2}\left(x_{i}\right) \\
& P(B) \geq 1-4(k t)^{-2}{ }_{i=k}^{n-k} \sigma^{2}\left(x_{i}\right)
\end{aligned}
$$

and the results follows.
3. Exact uniform confidence bounds for $\mathrm{E} \hat{\mathrm{m}}(\mathrm{x})$. Consistency.

Since $\inf _{x}[\hat{m}(x)-E \hat{m}(x)]=\min \left\{S_{i} ; i=0, \cdots, n-k\right\}$, them by the Lemma, $P\left(\inf _{x}[\hat{m}(x)-E \hat{m}(x)] \geqslant-t\right) \geq 1-4(k t)^{-2}\left[\sum_{i=1}^{n-k} \sigma^{2}\left(x_{i}\right)+{ }_{i=k}^{n} \sigma^{2}\left(x_{i}\right)\right]$.

Thus if we set

$$
\left.t_{\alpha}=\left(2 / \alpha^{\frac{1}{2}} k\right) \sum_{i=1}^{n} \bar{E}_{1}^{k} \sigma^{2}\left(x_{i}\right)+\sum_{i=k}^{n} \sigma^{2}\left(x_{i}\right)\right]^{\frac{1}{2}}
$$

then $\hat{m}(x)+t_{\alpha}$ is a simultaneous upper confidence boundary for $E \hat{m}(x)$ with confidence coefficient at least $1-\alpha$.

Similarly, $\hat{m}(x)-t_{\alpha}$ is a simultaneous lower confidence boundary for $\hat{E} \hat{m}(x)$ with confidence coefficient at least $1-\alpha$. Let $T^{ \pm}=\sup _{x} \pm[\hat{m}(x)-E \hat{m}(x)]$ and $T=\max \left\{T, T^{+}\right\}$then $P(T \geqslant t) \leq P\left(T^{+} \geqslant t\right)+P\left(T^{-} \leqslant-t\right)$. It follows that, for each $t>0$,

$$
P\left(\sup _{x}|\hat{m}(x)-E \hat{m}(x)| \leqslant t\right) \geq 1-8(k t)^{-2}\left[\sum_{i=1}^{n} \sum_{1}^{-k} \sigma^{2}\left(x_{i}\right)+\sum_{i=k}^{n} \sigma^{2}\left(x_{i}\right)\right] .
$$

Thus $\hat{m}(x)$ is a uniformly consistent estimate of $E \hat{m}(x)$ provided $k^{-2} \sum_{i=1}^{n} \sigma^{2}\left(x_{j}\right)$ tends to zero as $n \rightarrow \infty$. If $\sigma^{2}(x)$ is bounded, this follows if $\left(n / k^{2}\right) \rightarrow 0$ as $n \rightarrow \infty$.

Note that

$$
\begin{equation*}
\hat{m}(x) \pm \sqrt{2} t_{\alpha} \tag{3.1}
\end{equation*}
$$

is a level $(1-\alpha)$ simultaneous confidence band for $E \hat{m}(x)$. If we assume that $\sigma^{2}(x) \equiv \sigma^{2}$ for all $x$, then

$$
t_{\alpha}=(2 \sqrt{2(n-k)+1} / \sqrt{\alpha} k) \sigma
$$

and the width of the confidence band is $(4 \sqrt{2} \sqrt{2(n-k)+1} / \sqrt{\alpha} k) \sigma$. If we choose $k=n^{\frac{1}{2}+\Delta}, \Delta \leq \frac{1}{2}$, then the width is

$$
8 n^{-\Delta} \sigma / \sqrt{\alpha}
$$

minus a smaller order term. Thus the width tends to zero (and we have consistency) provided $\frac{1}{2}<\Delta<1$.

To use the band, we need an estimate of $\sigma^{2}$. A natural estimate is the residual mean square

$$
\begin{equation*}
\hat{\sigma}^{2}=n^{-1} \sum_{i=1}^{n}\left[Y_{i}-\hat{m}\left(x_{i}\right)\right]^{2} \tag{3.2}
\end{equation*}
$$

4. Bounds based on asymptotic distribution theory.

We assume $\sigma^{2}\left(x_{i}\right) \equiv \sigma^{2}<\infty$. Since the $\widetilde{S}_{\mathbf{i}}, i=1,2, \ldots$ are partial sums of independent identically distributed random variables, we may apply a result for Brownian motion (e.g., Billingsley (1968), p. 72) which yields

$$
P(A)=P\left(\max _{0 \leqslant i \leqslant n-k}\left(-\widetilde{S}_{i+k}\right)<\frac{1}{2} t\right) \approx 2 \Phi\left(\frac{1}{2} t \cdot k /(\sqrt{n} \cdot \sigma)\right)-1 .
$$

Similarly we get $P(B) \approx 2 \Phi\left(\frac{1}{2} t \cdot k /(\sqrt{n} \cdot \sigma)\right)-1$. As in Section 3 we find that asymptotically a level ( $1-\alpha$ ) confidence band for $E \hat{m}(x)$ is

$$
\begin{equation*}
\hat{m}(x) \pm 2 \Phi^{-1}(1-\alpha / 4) \sigma \cdot n^{\frac{1}{2}} / k . \tag{4.1}
\end{equation*}
$$

Choosing $k=n^{\frac{1}{2}+\Delta}$, the width of the band is

$$
4 \Phi^{-1}\left(1-\frac{1}{4} \alpha\right) n^{-\Delta} \sigma
$$

The following table compares the widths of the confidence bands (3.1) and (4.1). The widths of the bands are of the form $2 c \sigma n^{-\Delta}$. Table 4.1. Values of $c$ for three frequently used confidence coefficients.

$$
\text { Values of } c
$$

| Band $\alpha$ | .1 | .05 | .01 |
| :--- | ---: | ---: | ---: |
| Exact (3.1) | 8.94 | 12.65 | 28.28 |
| Asymptotic (4.1) | 3.92 | 4.48 | 5.63 |

5. Bounds based on nonoverlapping neighborhoods.

The bands in Sections 3 and 4 will be of use only for large data sets. In this section, we develop a band which is much narrower, but it is simultaneous only for a sequence $t_{n 1}, \cdots, t_{n 1}$ of $x$-values. The model is

$$
y_{n i}=m\left(x_{n i}\right)+\epsilon_{n i}, \quad x_{1 n}<\cdots<x_{n n}
$$

where $\epsilon_{n 1}, \cdots, \epsilon_{n 1}$ are independent with $\operatorname{Var}\left(\epsilon_{n 1}\right)=\sigma^{2}\left(x_{n i}\right)$ and

$$
\max _{j}\left|x_{n, j+k}-x_{n j}\right|=0\left(n^{\Delta^{-\frac{1}{2}}}\right), \quad 0<\Delta<\frac{1}{2}
$$

If we choose $t_{n 1}, \cdots, t_{n 1}$ so that

$$
\max _{j}\left|t_{n, j+1}-t_{n j}\right|=0\left(n^{-\lambda}\right), \quad 0 \leq \lambda \leq \frac{1}{2}-\Delta
$$

then for $n$ large enough, there will be no overlap between the $k=n^{\frac{1}{2}+\Delta}$ nearest $x$-neighbors to the points $t_{n 1}, \cdots, t_{n 1}$. Thus if we define

$$
T_{n i}=\hat{m}\left(t_{n i}\right)-E \hat{m}\left(t_{n i}\right), \quad i=1, \cdots, 1
$$

then there exists $N$ such that $T_{n 1}, \cdots, T_{n 1}$ are independent for all $n \geqslant N$. By Chebychev's inequality and (2.4), for $n \geq N$, $a>0$,

$$
P\left(\max _{j}\left|T_{n i}\right| \leq a\right) \geq \prod_{i=1}^{1}\left[1-(a k)^{-2}{\underset{j}{i}+k}_{\sum_{i+1}} \sigma^{2}\left(x_{n i}\right)\right] .
$$

If we assume $\sigma^{2}(x) \equiv \sigma^{2}$ then

$$
\begin{equation*}
\hat{m}(x) \pm \sigma /\left\{k\left[1-(1-\alpha)^{\frac{1}{7}}\right\}^{\frac{1}{2}}\right. \tag{5.1}
\end{equation*}
$$

is a simultaneous confidence band for $E \hat{m}(x)$ valid for all $x \in\left\{t_{n 1}, \cdots, t_{n\}}\right\}$. The width of this band is of the order $0\left(n^{-\frac{1}{4}-\frac{1}{2} \Delta}\right)$. By (2.4) and the Central Limit Theorem,

$$
\lim _{n \rightarrow \infty} P\left(\sqrt{k} T_{n j} \leq t\right)=\Phi(t / \sigma),
$$

where $k=n^{\frac{1}{2}+\Delta}$, and 1 is finite. Thus if we set

$$
M_{1}=\max \left\{\left|T_{n 1}\right|, \cdots,\left|T_{n 1}\right|\right\}
$$

then

$$
\begin{equation*}
\lim _{n \rightarrow \infty} P\left(\sqrt{k} M_{1} \leq t\right)=\left[2 \Phi\left(\frac{\sqrt{k} t}{\sigma}\right)-1\right] \tag{5.2}
\end{equation*}
$$

and

$$
\begin{equation*}
\hat{m}(x) \pm \Phi^{-1}\left(\frac{1}{2}[1+(1-\alpha)]^{\frac{1}{1}}\right) \sigma / \sqrt{k} \tag{5.3}
\end{equation*}
$$

is asymptotically a level ( $1-\alpha$ ) simultaneous confidence band for $E \hat{m}(x)$ valid for $x \in\left\{t_{n 1}, \cdots, t_{n 1}\right\}$.

This band is of order $0\left(n^{-\frac{1}{4}-\frac{1}{2} \Delta}\right)$. Note that (5.1) and (5.3) are considerably narrower than the bands (3.1) and (4.1) .

We next derive approximations to (5.2) and (5.3) valid for large 1. Let $V_{1}=\max _{i}\left\{T_{n i}\right\}, W_{1}=\min _{i}\left\{T_{n i}\right\}$, then $M_{1}=\max \left\{V_{1},-W_{1}\right\}$ and

$$
P\left(\sqrt{k} M_{1} \leq t\right)=P\left(\sqrt{k} V_{1} \leq t,-\sqrt{k} W_{1} \leq t\right) .
$$

Using this and results on the asymptotic distribution of extreme order statistics (e.g. Galambos, p. 65 and p. 106), we find that if

$$
a_{1}=-\Phi^{-1}\left(\frac{1}{1}\right) \text { or } a_{1}=(2 \log 1)^{\frac{1}{2}}-\frac{\frac{1}{2}(\log \log 1+\log 4 \Pi)}{(2 \log 1)^{\frac{1}{2}}}
$$

and $b_{1}=1 /(2 \log 1)^{\frac{1}{2}}$, then

$$
\begin{equation*}
\lim _{1 \rightarrow \infty}\left[\lim _{n \rightarrow \infty} P\left(\frac{\sqrt{k}}{\sigma} M_{1} \leq a_{1}+b_{1} z\right)\right]=\exp \left(-2 e^{-z^{\prime}}\right) \tag{5.4}
\end{equation*}
$$

It follows that for large 1 , an approximation to (5.3) is

$$
\begin{equation*}
\hat{m}(x) \pm \sigma\left(a_{1}+b_{1} z_{\alpha}\right) / \sqrt{k} \tag{5.5}
\end{equation*}
$$

where $\quad z_{\alpha}=-\log \left[-\frac{1}{2} \log (1-\alpha)\right]$.
It would have been more elegant to take the limit in (5.4) as 1 and $n$ simultaneously tend to $\infty$, say by setting $1=n^{\gamma}, 0<\gamma \leq \frac{1}{2}-\Delta$. With $\gamma=\frac{1}{2}-\Delta$, this would lead to a band similar to that of Revesz (1979). His estimator differs slightly from $\hat{m}(x)$ in that his index sets, say $I_{n k}^{\prime}(x)$, are balanced with $\frac{1}{2} k$ values of $\dot{x}_{n 1}, \cdots, x_{n n}$ closest to and less than $x$, and $\frac{1}{2} k$ values closest to and greater than $x, k$ even. Nevertheless, Theorem 1 of Reversz (1979) holds for $\hat{m}(x)$ based on the sets $I_{n k}(x)$, yielding the confidence band

$$
\begin{equation*}
\hat{m}(x) \pm \sigma\left(a_{s}+b_{s} z_{\alpha}\right) / \sqrt{k} \tag{5.6}
\end{equation*}
$$

with $s=n^{\frac{1}{2}-\Delta}$. The width of this band is of order $0\left(n^{-\frac{1}{4}-\frac{1}{2} \Delta}(\log n)^{\frac{1}{2}}\right)$ when $k=n^{\frac{1}{2}+\Delta}$. Note that in (5.1), (5.3) and (5.5) as well as (3.1) and (4.1) we have avoided a number of regularity conditions required by Révész.

The band (5.1) can be made asymptotically valid for all $x$ provided the $t_{n j}$ 's are chosen dense in the set $x$ of possible $x$ 's, that the bias is of smaller order than the widths of the bands (see Section 6), and that $m(x)$ is uniformly continuous on X..
6. Bias.

The bands of the previous sections are for

$$
\bar{m}(x)=E \hat{m}(x)=\sum_{i \in I_{n k}(x)^{m\left(x_{i}\right) / k} .}
$$

In order to make them valid for $m(x)$, we need to show that the bias

$$
m(x)-\bar{m}(x)=\sum_{i \in I_{n k}(x)}^{\left[m\left(x_{i}\right)-m(x)\right] / k}
$$

is uniformly of smaller order than the width of the bands.
Assume now that $x$ is in an interval [a,b] and that the regression function satisfies a rth order Lipschitz condition:

$$
\begin{equation*}
|m(x)-m(y)| \leqslant c|x-y|^{r}, \quad x, y \in R . \tag{6.1}
\end{equation*}
$$

Then,

$$
\begin{aligned}
|\bar{m}(x)-m(x)| & \leqslant c \cdot \sum_{j \in I_{n k}(x)\left|x_{j}-x\right|^{r} / k} \\
& \leqslant c \cdot \max \left\{\left(\frac{x_{i+k^{-x}}}{2}\right)^{r}, \quad\left(\frac{x_{i+k+j^{-x}}}{2}\right)^{r}\right\} \\
& \leqslant c \cdot\left(x_{i+k+1}-x_{i}\right)^{r} / 2^{r},
\end{aligned}
$$

when $x \in J_{i}=\left(\left(x_{i}+x_{i+k}\right) / 2,\left(x_{i+1}+x_{i+k+1}\right) / 2\right], i=0,1, \ldots, n-k$, $x_{0}=a, \quad x_{n+1}=b$.

Thus when $k=n^{\frac{1}{2}+\Delta}$ and $\max _{j}\left|x_{n, j+k+1}^{-x_{n j}}\right|=0\left(n^{\Delta-\frac{1}{2}}\right), 0<\Delta<\frac{1}{2}$, where now $x_{i}$ and $x_{n i}$ are used interchangeably, then

$$
\begin{equation*}
\sup _{x}|\bar{m}(x)-m(x)|=0\left(n^{-r\left(\frac{1}{2}-\Delta\right)}\right) \tag{6.2}
\end{equation*}
$$

and the bias is of smaller order than the width of the bands (3.1) and (4.1) when $n^{-\Delta}>n^{-r\left(\frac{1}{2}-\Delta\right)}$, i.e. $\Delta<r / 2(r+1)$. Thus, if we are only
willing to assume a first order Lipschitz condition, we must chose $\Delta$ less than $\frac{1}{4}$. Second and third order conditions lead to $\Delta$ less than $\frac{1}{3}$ and $\frac{3}{8}$, respectively.

Turning to the bands (5.1) and (5.3), we need $n^{-\frac{3}{4}-\frac{1}{2} \Delta}>n^{-r\left(\frac{1}{2}-\Delta\right)}$, i.e. $\Delta<\left(\frac{1}{2} r-\frac{1}{4}\right)\left(\frac{1}{2}+r\right)^{-1}$. For $r=1,2$ and 3 , we find $\Delta$ less than $\frac{1}{6}, \frac{3}{10}$ and $\frac{5}{14}$, respectively. Since the bands (5.5) and (5.6) are at least as wide as (5.1) and (5.3), the same restriction is sufficient to make the bias of uniformly smaller order than the widths.

Revesz (1979) considers a different restriction of the regressor to the interval $[0,1]$ : Let $X$ have a density $f$ such that $f(x) \geqslant \lambda$, $x \in[0,1]$, for some $\lambda>0$. Let $k$ be such that

$$
\begin{equation*}
\mathrm{kn}^{-2 / 3} \log n \rightarrow 0, \quad k^{-1}(\log n)^{3} \rightarrow 0 \text { as } n \rightarrow \infty \tag{6.3}
\end{equation*}
$$

Then (RÉvesz (1979), Lemma 1)

$$
\lim _{n \rightarrow \infty} \sup \frac{n}{k} \sup _{1 \leqslant i \leqslant n-k}\left|x_{i}-x_{i+k}\right| \leqslant 2 / \lambda \quad \text { a.s. }
$$

From this and the rth order Lipschitz condition (6.1), it follows that if $k=n^{\frac{1}{2}+\Delta}$,

$$
\sup _{x}|\bar{m}(x)-m(x)|=0\left(n^{-r\left(\frac{1}{2}-\Delta\right)}\right) \text { a.s. }
$$

This result is valid for $X \in[a, b]$, not just $[0,1]$, thus we have shown (6.2) again under a different set of conditions. Note that (6.3) implies that if $k=n^{\frac{1}{2}+\Delta}$, then $\Delta<\frac{1}{6}$.

Under the conditions of the above paragraph, and assuming that $m(x)$ has a uniformly bounded derivative, Révesz has shown (Lemma 2) that

$$
\sup _{x}|\bar{m}(x)-m(x)|=o\left((k \log n)^{-\frac{1}{2}}\right) \text { a.s. }
$$

which again implies that the bias is of smaller order than the width of the band.

Rosenblatt (1969) gives asymptotic results for kernel estimators of the regression function. He gets pointwise confidence-intervals of width $n^{-2 / 5}$ using a bandwidth of order $n^{-1 / 5}$ which corresponds to $k=n^{4 / 5}$ or $\Delta=.3$.

Spiegelman and Sacks (1980) obtain window estimators with mean squared error of the order $0\left(n^{-2 / 3}\right)$ by imposing a first order Lipschitz condition on $m$. Thus their bias is of order less than $n^{-1 / 3}$. Their bandwidth is of order $b_{n}=n^{-1 / 3}$, corresponding to $k=n^{2 / 3}$.
7. An illustration.

To get an idea of the accuracy of the bands, we computed the band (5.6) for data $\left(x_{1}, Y_{1}\right) \cdots\left(x_{100}, Y_{100}\right)$ generated from the model

$$
\begin{equation*}
y_{i}=e^{\beta_{1}} e^{\beta_{2} x_{i}} x_{i}^{\beta_{3}}+e_{i} \tag{7.1}
\end{equation*}
$$

where $\beta_{1}=5, \quad \beta_{2}=-\frac{1}{2}, \quad \beta_{3}=1, e_{1}, \cdots, e_{n}$ are i.i.d. $N(0,100)$, $x_{i}=i / 25, \quad i=1, \cdots, 100$, the significance coefficient is $.90, \hat{\sigma}$ is computed from (3.2), and $\Delta=.15$.

This model has been suggested for agricultural experiments where an amount $x$ of fertilizer increases yield $Y$ for low and moderate doses while it decreases yield at high doses.

The result is shown in Figure 1 where the middle curve is the estimate $\hat{m}(x)$ and the upper and lower curves define the band. The band is fairly accurate with the width being 11.5. Since the band is simultaneous, we can test model assumptions. For instance, since no line fits in the confidence band, regression linear in $x$ is rejected. Similarly, a parabola does not fit in the band and a quadratic (in $x$ ) regression model is also rejected.

We also computed the widths of the other bands (using $\sigma=10$ rather than $\hat{\sigma}$ ). The results are given in the table below using $k=20$ and $1=s=5$. Note that the widths of (5.5) and (5.6) are the same.

Table 7.1. Widths of the confidence bands for model (7.1).

| Band | $(3.1)$ | $(4.1)$ | $(5.1)$ | $(5.3)$ | $(5.6)$ |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Width | 89.6 | 39.3 | 31.0 | 10.3 | 11.6 |

Note that since $e_{i}$ is normal, (5.3) is exactly a level . 90 simultaneous confidence procedure for $1=5$.


Figure 1. The estimate $\hat{m}(x)$ and $90 \%$ simultaneous confidence band for the model (7.1).

## Acknowl edgement

We wish to thank Professor J. Mac Queen for generously spending his time discussing the topic of this paper.

## REFERENCES

[1] Billingsley, P. (1968). Convergence of Probability Measures. Wiley, New York.
[2] Bhattacharya, P. K. (1974). Convergence of sample paths of normalized sums of induced order statistics. Ann. Statist. 2, 1034-1039.
[3] Loève, M. (1963). Probability Theory, Third Edition. Academic Press, New York.
[4] Revesz, P. (1979). On the nonparametric estimation of the regression function. Problems of Control and Info. Theo. 8, 297-302.
[5] Rosenblatt, M. (1969). Conditional probability density and regression estimates. International Symposium on Multivariate Analysis, II.
[6] Spiegelman, C. and Sacks, J. (1980). Consistent window estimation in nonparametric regression. Ann. Statist. 8, 240-246.
[7] Stone, C. (1977). Consistent nonparametric regression. Ann. Statist. 5, 595-620.

```
Steinar Bjerve
Matematisk Institutt
P.0. Box 1053, Blindern
Oslo 3, Norway
Kjell A. Doksum
Department of Statistics
University of California
Berkeley, CA }9472
U.S.A.
Brian S. Yandell
Statistics Department
University of Wisconsin
1210 West Dayton St.,
Madison, WI 53706
U.S.A.
```


[^0]:    $1_{\text {This }}$ work was done while the author was visiting the University of California at Berkeley on sabbatical leave from the University of Oslo.
    ${ }^{2}$ Doksum's work was partially supported by National Science Foundation Grant MCS-81-02349.

