NONPARAMETRIC ESTIMATION OF A REGRESSION FUNCTION

by

Kuang-Fu Cheng and Pi-Erh Lin

FSU Statistics Report No. M512
ONR Technical Report No. 145

July, 1979
The Florida State University
Department of Statistics
Tallahassee, Florida 32306

1Research supported by the Army, Navy and Air Force under Office of Naval Research Contract No. N00014-76-C-0608. Reproduction in whole or in part is permitted for any purpose of the United States Government.
Nonparametric Estimation of a Regression Function

by

Kuang-Fu Cheng and Pi-Erh Lin
The Florida State University

Summary

Consider the regression model \( Y_i = g(X_i) + e_i, \ i = 1, \ldots, n, \) where \( g \) is an unknown function defined on the interval \([0, 1]\). Fix \( 0 < x_1 \leq x_2 \leq \ldots \leq x_n \leq 1, \) with \( \max(x_1 - x_{i-1}) = o(n^{-1}) \), and assume \( \{e_i\} \) are iid random variables with zero mean and finite variance. Priestley and Chao (J. Roy. Statist. Soc. B 34 (1972), 385-392) and Clark (J. Roy. Statist. Soc. B 39 (1977), 107-113) proposed estimators \( g_{2n} \) and \( g_{3n} \), respectively, for \( g \). In this paper, asymptotic behaviors of \( g_{2n} \) and \( g_{3n} \) are studied utilizing the properties of the new estimator \( g_{1n}(x) = \prod_{i=1}^{n} Y_i \{K[(x - x_{i-1})/a_n] - K[(x - x_i)/a_n]\} \) where \( K(u) \) is a known absolutely continuous cdf with pdf \( k(u) \). It is shown that \( g_{1n}, g_{2n}, \) and \( g_{3n} \) are asymptotically equivalent in various senses. Moreover, consistent results are established and rates of uniform convergence obtained. For example, if \( E|Y_i|^3 < \infty \), if both \( g \) and \( k \) are Lipschitz of order 1, and if \( \{\beta_n\} \) is any diverging sequence as \( n \to \infty \), then, for \( 0 < a \leq b < 1, \)

\[
(n^{1/3}/\beta_n \log n) \sup_{a \leq x \leq b} [g_{1n}(x) - g(x)] \overset{w.p.1}{\to} 0, \text{ as } n \to \infty.
\]

Finally, when \( g \) is monotone, a strong consistent isotonic estimator \( \hat{g}_n \) is considered.

Research supported by the Army, Navy, and Air Force under Office of Naval Research Contract No. N00014-76-C-0608. Reproduction in whole or in part is permitted for any purpose of the United States Government.

AMS 1970 subject classifications: Primary 62G05; Secondary 60F15.

Key words and phrases: Convergence rate; isotonic; Lipschitz; strongly uniform convergence.
1. **Introduction.** Let \( Y_1, \ldots, Y_n \) be \( n \) independent observations at fixed \( x_1, \ldots, x_n \) according to the nonlinear regression model

\[
Y_i = g(x_i) + e_i
\]

where \( g(x) \) is an unknown function defined on the closed interval \([0, 1]\) and \( \{e_i\} \) are i.i.d. random variables with zero mean and finite variance \( \sigma^2 \). The problem of estimating \( g(x) \) was first considered by Priestley and Chao (1972) and later studied by Benedetti (1977), Clark (1977), and Schuster and Yakowitz (1979). Since the design variables \( x_1, \ldots, x_n \) are selected and fixed in \([0, 1]\), we may assume without loss of generality that \( 0 = x_0 \leq x_1 \leq \ldots \leq x_n \leq 1 \) and define \( \delta_n = \max_{1 \leq i \leq n} (x_i - x_{i-1}) \). The design scheme requires \( \delta_n \) to be sufficiently small when \( n \) is sufficiently large. This requirement is not needed in the study of a different nonlinear regression model such as \( m(x) = E(Y|X = x) \). For more detail on the study of this model see, e.g., Nadaraya (1964), Watson (1964), Schuster (1972), Ahmad and Lin (1976), and Cheng (1979), among others.

In the next section we study the properties of consistency and obtain rates of convergence for the estimators of Priestley and Chao (1972) and Clark (1977). In so doing, we find it useful to propose a new estimator based on a "naive" initial estimator. The consistency results for the proposed estimator are studied in detail and those for the other two estimators are obtained indirectly from a lemma relating the three estimators in an asymptotic sense. Finally, in Section 3, when \( g(x) \) is monotone, an isotonic estimator is considered and shown to be strongly consistent.

For the sake of comparison and completeness we state the three estimators under consideration.
(a) The proposed estimator is given by

\[ g_{1n}(x) = \sum_{i=1}^{n} Y_i \int_{x_{i-1}}^{x_i} a_n^{-1} k[(x - z)/a_n] \, dz \]

where \( k(z) \) is a kernel function satisfying (i) \( k(z) \geq 0 \) for all \( z \), (ii) \( k(z) = 0 \) for \( z \notin [-L, L] \) for some positive constant \( L \), and (iii) \( \int_{-L}^{L} k(z) \, dz = 1 \), and where \( \{a_n\} \) is a sequence of positive constants converging to \( 0 \) as \( n \to \infty \). The estimator \( g_{1n}(x) \) is obtained by the initial "naive" estimator

\[ h_1(x) = \begin{cases} 
Y_1 & \text{for } x \leq x_1 \\
Y_{i+1} & \text{for } x_i < x \leq x_{i+1}, \ i = 1, \ldots, n-1 \\
Y_n & \text{for } x \geq x_n 
\end{cases} \]

smoothing with the weight function \( a_n^{-1} k[(x - z)/a_n] \). That is, for all \( x \in (0, 1) \) and all \( n \) sufficiently large

\[ g_{1n}(x) = \int_{-\infty}^{\infty} a_n^{-1} k[(x - z)/a_n] h_1(z) \, dz. \]

(b) Clark (1977) proposed the following piecewise-linear continuous function as an initial estimator

\[ h_2(x) = \begin{cases} 
Y_1 & \text{for } x \leq x_1 \\
Y_i \left( \frac{x_{i+1} - x}{x_{i+1} - x_i} \right) + Y_{i+1} \left( \frac{x - x_i}{x_{i+1} - x_i} \right) & \text{for } x_i < x \leq x_{i+1} \\
Y_n & \text{for } x \geq x_n, \ i = 1, \ldots, n-1 
\end{cases} \]

and then smoothed it with the same weight function \( a_n^{-1} k[(x - z)/a_n] \) to obtain

\[ g_{2n}(x) = \sum_{i=1}^{n} c_{ni} Y_i \]
where, for all $x \in (0,1)$ and all $n$ sufficiently large,

$$ (1.6) \quad c_{ni} = \begin{cases} 
\int_{x_0}^{x_1} a^{-1} k(\frac{x-z}{a_n})dz + \int_{x_1}^{x_2} a^{-1} k(\frac{x-z}{a_n})dz, & i = 1 \\
\int_{x_{i-1}}^{x_i} a^{-1} k(\frac{x-z}{a_n})dz + \int_{x_i}^{x_{i+1}} a^{-1} k(\frac{x-z}{a_n})dz & i = 2, \ldots, n-1 \\
\int_{x_{n-1}}^{x_n} a^{-1} k(\frac{x-z}{a_n})dz, & i = n.
\end{cases} $$

(c) The third estimator of $g(x)$ was introduced by Priestley and Chao (1972) and is given by

$$ (1.7) \quad g_{3n}(x) = \sum_{i=1}^{n} \gamma_i (x_i - x_{i-1}) a^{-1} k(\frac{x-x_i}{a_n}). $$

Mathematically speaking, if $\delta_n$ is sufficiently small these three estimators are asymptotically equivalent. In the next section we will present detailed argument about this fact.

2. Properties of Consistency and Rates of Convergence. For convenience of presentation the following notations will be adopted. Define

(I) $A_i = \{ x : x_i \leq x \leq x_{i+1} \}, \ i = 0, \ldots, n-1;$

(II) $B_i = B_i(x) = \int_{A_i} \frac{x_{i+1} - z}{a_n} a^{-1} k(\frac{x-z}{a_n})dz;$

(III) "$f(x) \in \text{Lip}(\alpha)$" means that $f(x)$ is a Lipschitz function of order $\alpha$, $\alpha > 0$; and

(IV) $\|f\|_{[a,b]} = \sup_{a \leq x \leq b} |f(x)|,$ $-\infty < a \leq b < \infty.$
The following lemma relates the interrelationship among estimators $g_{1n}(x)$, $g_{2n}(x)$, and $g_{3n}(x)$ under various conditions on the regression function $g(x)$ and the kernel function $k(z)$. It also provides the rates of convergence of the mean square error for $g_{in}(x)$, $i = 1, 2, 3$.

**Lemma 1.** (i) If $g(x) \in \text{Lip}(a)$ and $||x||_{[-L,L]} < \infty$, then

$$||E(g_{1n} - g_{2n})^2||_{[a,b]} = O(\delta_n a_n^{-1} + \delta_n^{2a})$$

for all $0 < a \leq b < 1$.

(ii) If $||g||_{[0,1]} < \infty$ and $k(z) \in \text{Lip}(\varepsilon)$, then

$$||E(g_{1n} - g_{3n})^2||_{[a,b]} = O(\delta_n^{2\beta+1}(1 + n^2 \delta_n)/a_n^{2\beta+2})$$

for all $0 < a \leq b < 1$.

**Proof.** The difference between the initial estimators may be expressed as

$$h_1(x) - h_2(x) = \begin{cases} 
\frac{(Y_{i+1} - Y_i)(x_{i+1} - x)}{x_{i+1} - x_i} & \text{if } x_i \leq x \leq x_{i+1} \\
0 & \text{if } x \leq x_1 \text{ or } x \geq x_n.
\end{cases}$$

Then, with an application of the $c_r$-inequality of Loève (1963, p. 155) it follows that, for $n$ sufficiently large,

$$E[(g_{1n}(x) - g_{2n}(x))^2] = E(\int_{-\infty}^{\infty} [h_1(z) - h_2(z)]a_n^{-1}k(\frac{x - z}{a_n})dz)^2.$$ 

$$\leq 2E\left(\sum_{i=1}^{n-1} [(Y_{i+1} - g(x_{i+1})) - (Y_i - g(x_i))]B_i^2 + 2\sum_{i=1}^{n-1} [g(x_{i+1}) - g(x_i)]B_i \right)^2$$

$$= 2(T_{1n} + T_{2n})$$ say.
It is clear that

\[ T_{1n} = \text{Var}\{ \sum_{i=1}^{n-1} (Y_{i+1} - Y_i)B_i \} \]

\[ = \sum_{i=1}^{n-1} 2\sigma^2 B_i  + \sum_{i=1}^{n-1} \text{cov}[ (Y_{i+1} - Y_i)B_i, (Y_{i+2} - Y_{i+1})B_{i+1} ] \]

\[ + \sum_{i=1}^{n-1} \text{cov}[ (Y_{i+1} - Y_i)B_i, (Y_i - Y_{i-1})B_{i-1} ] \]

\[ = \sigma^2 \sum_{i=1}^{n-1} (2B_i - B_{i+1}B_i - B_{i-1}B_i) \cdot \]

Note that, for \( i = 1, \ldots, n - 1 \),

\[ 0 \leq B_i \leq \int_{A_i} a_n^{-1} k(\frac{x-z}{a_n})\,dz \leq \delta a_n^{-1} ||k||_{[-L,L]} \]

and that \( 0 \leq \sum_{i=1}^{n-1} B_i \leq 1 \). Consequently,

\[ |T_{1n}| \leq 4\sigma^2 \delta a_n^{-1} ||k||_{[-L,L]} \sum_{i=1}^{n-1} B_i \]

\[ = O(\delta a_n^{-1}) \cdot \]

As for \( T_{2n} \), since \( g(x) \in \text{Lip}(\alpha) \), we have

\[ T_{2n} = \left( \sum_{i=1}^{n-1} [g(x_{i+1}) - g(x_i)]B_i \right)^2 \]

\[ \leq \delta^{2\alpha} A \cdot M \cdot \]

where \( 0 < M < \infty \). This proves assertion (i).
Regarding (ii), we consider \( E[g_{1n}(x) - g_{3n}(x)]^2 \) which can be expressed as the sum of the variance and mean square. That is, for all \( x \in (0, 1) \),

\[
E[g_{1n}(x) - g_{3n}(x)]^2 = E\left( \sum_{i=1}^{n} Y_i \int_{x_{i-1}}^{x_i} a_n^{-1} [k(\frac{x-x_i}{a}) - k(\frac{x-z}{a})] dz \right)^2
\]

\[
= \sigma^2 \sum_{i=1}^{n} (\int_{x_{i-1}}^{x_i} a_n^{-1} [k(\frac{x-x_i}{a}) - k(\frac{x-z}{a})] dz)^2
\]

\[
+ \left( \sum_{i=1}^{n} g^2(x_i) \int_{x_{i-1}}^{x_i} a_n^{-1} [k(\frac{x-x_i}{a}) - k(\frac{x-z}{a})] dz \right)^2
\]

\[
= C \sigma^2 \sum_{i=1}^{n} (\frac{\delta}{a_n})^{2-\beta} (\frac{\delta}{a_n})^{2-\beta} \int_{x_{i-1}}^{x_i} dz + o(n^2 (\frac{\delta}{a_n})^{2-\beta+2}), \quad 0 < \beta < \infty,
\]

\[
= O(\frac{\delta^{2\beta+1}}{a_n^{2\beta+2}} + o(n^2 \frac{\delta^{2\beta+2}}{a_n^{2\beta+2}}),
\]

completing the proof of (ii). □

The following theorem establishes the pointwise weak consistency for \( g_{1n}(x), i = 1, 2, 3 \). It is noted that Priestley and Chao (1972) obtained the weak consistency of \( g_{3n}(x) \) under a set of conditions which are stronger than those established below.

**Theorem 1**  
(i) Assume that \( k(z) \) is a bounded function and \( \frac{\delta}{a_n} \rightarrow 0 \) as \( n \rightarrow \infty \). If \( g(x) \) is continuous in \( [0, 1] \), then \( g_{1n}(x) \xrightarrow{P} g(x) \) as \( n \rightarrow \infty \) for all \( x \in (0, 1) \).

(ii) Assume the conditions of (i). If \( g(x) \in \text{Lip}(\alpha) \), then \( g_{2n}(x) \xrightarrow{P} g(x) \) as \( n \rightarrow \infty \) for all \( x \in (0, 1) \).
(iii) Assume that \( k(z) \in \text{Lip}(\beta), \delta_n = O(n^{-1}) \), and that \( na_n^{1+1/\beta} \to \infty \) as \( n \to \infty \). If \( g(x) \) is continuous in \([0, 1]\), then \( g_{3n}(x) \overset{P}{\to} g(x) \) as \( n \to \infty \) for all \( x \in (0, 1) \).

**Proof.** The estimator \( g_{1n}(x) \) can be regarded as a weighted sum of independent random variables: Thus utilizing Theorem 1 of Pruitt (1966) we obtain

\[
g_{1n}(x) - Eg_{1n}(x) \overset{P}{\to} 0, \text{ as } n \to \infty.
\]

Furthermore, for any \( x \in (0, 1) \), the bias is

\[
Eg_{1n}(x) - g(x) = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} [g(x_i) - g(z)] a_n^{-1} k\left(\frac{x - z}{a_n}\right) dz
\]

\[
+ \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} [g(z) - g(x)] a_n^{-1} k\left(\frac{x - z}{a_n}\right) dz, \text{ for large } n.
\]

Thus, by the uniform continuity of \( g(x) \) in \([0, 1]\),

\[
(1.8) \quad Eg_{1n}(x) - g(x) \to 0 \text{ as } n \to \infty.
\]

That is, \( g_{1n}(x) \) is asymptotically unbiased. Assertion (i) is now proved by the triangular inequality. Conclusions of (ii) and (iii) follow from (i) and Lemma 1. \( \square \)

The next theorem is a refinement of Theorem 1. It obtains the convergence rates for the uniform mean square error for \( g_{in}(x), i = 1, 2, 3 \).
THEOREM 2. (i) Assume that $k(z)$ is a bounded function. If $g(x) \in \text{Lip}(\alpha)$, then for $0 < a \leq b < 1$

$$||E(g_{in} - g)^2||_{[a,b]} = O(\delta_n a^{-1} n^{-1} + \delta_n^{2\alpha} + a_n^{2\alpha}), i = 1, 2.$$  

(ii) If $k(x) \in \text{Lip}(\beta)$ and $||g||_{[0,1]} < \infty$, then for $0 < a \leq b < 1$

$$||E(g_{3n} - g)^2||_{[a,b]} = O(\delta_n a^{-1} n^{-1} + \delta_n^{2\alpha} + a_n^{2\alpha} + \delta_n^{2\beta+1}(1 + n^2 \delta_n) / a_n^{2\beta+2}).$$

PROOF. We will only prove the uniform mean square consistency and obtain the convergence rate for $g_{1n}(x)$. Note that

$$E[g_{1n}(x) - g(x)]^2$$

$$= \text{Var}[g_{1n}(x)] + [Eg_{1n}(x) - g(x)]^2$$

$$\leq \sigma^2 \sum_{i=1}^{n} \left[ \int_{x_{i-1}}^{x_i} a_n^{-1} k(\frac{x - z}{a_n}) dz \right]^2$$

$$+ 2\left( \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} [g(x_i) - g(z)] a_n^{-1} k(\frac{x - z}{a_n}) dz \right)^2$$

$$+ 2\left( \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} [g(z) - g(x)] a_n^{-1} k(\frac{x - z}{a_n}) dz \right)^2,$$

for large $n,$

$$= O(\delta_n a^{-1} n^{-1} + \delta_n^{2\alpha} + a_n^{2\alpha})$$

for large $n$ and for all $x \in [a, b]$.

This completes the proof of (i) for $i = 1$. Conclusion of (i) for $i = 2$ and (ii) follow from (i) for $i = 1$ in conjunction with Lemma 1. □

REMARKS. Some special choice of $a_n$ and $\delta_n$ are of interest. Consider, for example,
(a) \( a_n = n^{-1/3} \) and \( \delta_n = 0(n^{-1}) \). If \( \alpha = \beta = 1 \), then

\[ ||E(g_{in} - g)^2||_{[a,b]} = O(n^{-2/3}), \quad i = 1, 2, 3. \]

(b) If \( g(x) \) has a bounded second derivative (implying \( g(x) \in \text{Lip}(1) \)) and if \( \int z k(z)dz = 0 \), then \( a_n^{2\alpha} \) can be replaced by \( a_n^4 \) in the conclusion of Theorem 2. In particular, if \( a_n = n^{-1/5} \), \( \beta = 1 \), and \( \delta_n = 0(n^{-1}) \), then

\[ ||E(g_{in} - g)^2||_{[a,b]} = O(n^{-4/5}), \quad i = 1, 2, 3. \]

This result is comparable with (4.5) of Priestley and Chao (1972) and also comparable with (3) of Benedetti (1977).

In the previous theorems we have seen that, under suitable conditions, \( g_{in}(x) \) for \( i = 1, 2, 3 \), are asymptotically equivalent in the sense of convergence in probability and in \( L_2 \). In the following, we will develop almost sure behavior for the estimator \( g_{in}(x) \). It can be shown, under proper premium and same technique, that \( g_{2n}(x) \) and \( g_{3n}(x) \) have the same almost sure behavior as that of \( g_{in}(x) \).

**THEOREM 3.** Assume that \( k(z) \) is a bounded function. If \( g(x) \) is continuous in \( [0, 1] \), \( E|Y|^{1+1/\gamma} < \infty \), and \( \delta_n a_n^{-1} = O(n^{-\gamma}) \) for some \( \gamma > 0 \), then for all \( x \in (0, 1) \)

\[ g_{in}(x) \xrightarrow{w.p.1} g(x) \text{ as } n \to \infty. \]

**PROOF.** Since \( g_{in}(x) \) is asymptotically unbiased, it suffices to show that, for all \( x \in (0, 1) \),

\[ g_{in}(x) - Eg_{in}(x) \xrightarrow{w.p.1} 0 \text{ as } n \to \infty. \]

But this is an immediate consequence of Theorem 2 of Pruitt (1966). \( \square \)
It is noted that Benedetti (1977, Theorem 3) showed
\[ g_{3n}(x) \overset{w.p.1}{\rightarrow} g(x) \text{ as } n \to \infty, \]
for all \( x \in (0, 1) \), assuming \( \text{E} e^4 < \infty \) and other regularity conditions on \( k(z), \delta_n, \) and \( a_n \).

Attention now is directed to establish the uniform strong consistency of the estimator \( g_{1n}(x) \), and the rate of this consistency. The technique used here is similar to that of Cheng (1979).

**Theorem 4.** Let \( \gamma \) be a positive constant and \( \{ \beta_n \} \) be a sequence of constants diverging to \( \infty \) as \( n \to \infty \). Suppose that \( g(x) \in \text{Lip}(\alpha) \), \( k(z) \in \text{Lip}(\beta) \), and \( \text{E}|Y|^p < \infty, \ p > 1 \). Assume that, for \( n \) sufficiently large, the following conditions are satisfied.

(i) \( (n^{\gamma}/\beta_n \log n)(\delta_n^{\alpha} + a_n^{\alpha}) = o(1) \);
(ii) \( n^{2\gamma}\delta_n^{\beta}/a_n^{\beta}\beta_n = o(1) \);
(iii) \( \delta_n^{\beta}/(n^{2\beta-1-\gamma-1/p}\beta_n^{\beta+1}\log n) \to 0 \) as \( n \to \infty \); and
(iv) \( n^{\gamma+1/p}\delta_n^{\beta}/a_n^{1/2} \leq (2||k||_{[-L,L]}^{-1} \) for large \( n \).

Then, for all \( 0 < a < b < 1 \),
\[ (n^{\gamma}/\beta_n \log n)||g_{1n} - g||_{[a,b]} \overset{w.p.1}{\rightarrow} 0 \text{ as } n \to \infty. \]

**Proof.** It is clear that
\[ ||Eg_{1n} - g||_{[a,b]} = O(\delta_n^{\alpha} + a_n^{\alpha}). \]

This, together with condition (i), implies that we need only show that
\begin{equation}
\left( n^\gamma/\beta_n \log n \right) | | g_{1n} - E g_{1n} | |_{[a,b]} \xrightarrow{w.p.1} 0 \text{ as } n \to \infty.
\end{equation}

To this end, recall the model given by (1.1) and define the truncated random variable

\begin{equation}
\bar{e}_i = e_i I(|e_i| \leq i^{1/p}), \ i = 1, \ldots, n,
\end{equation}

where \( I(\cdot) \) is the usual indicator function. Using (2.2) we define an auxiliary variable as

\begin{equation}
\bar{g}_n(x) = \sum_{i=1}^{n} \bar{e}_i \int_{x_{i-1}}^{x_i} a_n^{-1} k(x-z)dz
\end{equation}

= \sum_{i=1}^{n} \bar{g}_n(i), \ \text{say}.

Then, for any \( \varepsilon > 0 \),

\begin{equation}
P\left( n^\gamma/\beta_n \log n | | \bar{g}_n(x) - E \bar{g}_n(x) | | > \varepsilon \right)
\end{equation}

\begin{align*}
&\leq \exp(\log n^{-\varepsilon \beta_n} 1/2) \prod_{i=1}^{n} \exp\left(n^\gamma/\beta_n 1/2\right) | | \bar{g}_n(i) - E \bar{g}_n(i) | |

&\leq n^{-\varepsilon \beta_n} 1/2 \prod_{i=1}^{n} \exp\left(n^2\gamma/\beta_n \right) \text{Var}[\bar{g}_n(i)],
\end{align*}

by an application of a moment inequality of the exponential form (see, e.g., Lamperti (1966), pp. 43-44). But

\begin{align*}
\sum_{i=1}^{n} \text{Var}[\bar{g}_n(i)]
\leq \sigma^2 \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} a_n^{-1} k(x-z)dz^2

= o(\varepsilon a_n^{-1}), \text{ for large } n \text{ and for all } x \in (0, 1).
\end{align*}
Thus, in view of condition (ii), we have

\[(2.4) \quad P\{ (n^{\gamma/\beta_n } \log n) | \bar{g}_n(x) - E\bar{g}_n(x) | > \varepsilon \} \leq C \cdot \varepsilon^{1/2} \]

\[\leq C \cdot n^{-\varepsilon \beta_n^{1/2}} \text{ for large } n \text{ and for all } x \in [a, b], \ 0 < C < \infty, \]

and, therefore,

\[(2.5) \quad P\{ (n^{\gamma/\beta_n } \log n) | \bar{g}_n(x) - E\bar{g}_n(x) | > \varepsilon \} = 0(n^{-\varepsilon \beta_n^{1/2}}) \]

for all \( x \in [a, b] \).

Next, we claim that

\[(2.6) \quad (n^{\gamma/\beta_n } \log n) | \bar{g}_n - E\bar{g}_n | |_{[a,b]} \xrightarrow{w.p.1} \text{ as } n \to \infty. \]

To show this, consider the set

\[D_n = \{x: \ |x| \leq n^{1/\delta} + L, \ x \in \mathbb{R} \}, \text{ for some } \delta > 0. \]

Let \( E_n \) be a set such that, for all \( x \in D_n \) there corresponds a point \( z(x) \in E_n \) satisfying \( |x - z(x)| < n^{-2} \) and such that \( E_n \) has at most \( N_n = [2n^2(n^{1/\delta} + L)] + 1 \) elements, where \([a] \) denotes the integer part of \( a \). Write

\[(2.7) \quad (n^{\gamma/\beta_n } \log n) | \bar{g}_n - E\bar{g}_n | |_{[a,b]} \cap D_n \leq S_{1n} + S_{2n} + S_{3n}, \]

where

\[S_{1n} = (n^{\gamma/\beta_n } \log n) \sup_{x \in [a,b] \cap D_n} | \bar{g}_n(x) - \bar{g}_n(z(x)) |, \]

\[S_{2n} = (n^{\gamma/\beta_n } \log n) \sup_{x \in [a,b] \cap D_n} | \bar{g}_n(z(x)) - E\bar{g}_n(z(x)) |, \]

and

\[S_{3n} = (n^{\gamma/\beta_n } \log n) \sup_{x \in [a,b] \cap D_n} | E\bar{g}_n(z(x)) - E\bar{g}_n(x) |. \]
Now \( P(S_{2n} > c) \leq \frac{C \cdot n^{-\frac{1}{2}}}{n} \) implies that \( S_{2n} \to 0 \) w.p.1, as \( n \to \infty \).

Moreover, we have

\[
S_{1n} \xrightarrow{\text{w.p.1}} (n^{\beta} \log n)^{1/p} \cdot (n^{2a-n})^{-\beta} \cdot \delta \cdot n^{-1} \cdot n
\]

\[ \to 0, \text{ as } n \to \infty, \]

by condition (iii). Similarly, \( S_{3n} \to 0 \) as \( n \to \infty \). Hence, the left hand side of (2.7) converges to 0, w.p.1, as \( n \to \infty \). But, for \( n \) sufficiently large, \([a, b] \cap D_n = [a, b]\). Assertion (2.6) follows. We are now in a position to prove (2.1). Define

(2.8) \( g_n(x) = \sum_{i=1}^{n} e_i \int_{x_{i-1}}^{x_i} a_n^{-1} k\left(\frac{x-z}{a_n}\right) dz, \)

and consider the inequality

(2.9) \( ||g_{1n} - E g_{1n}||_{[a,b]} = ||g_n^* - E g_n^*||_{[a,b]} \)

\[ \leq ||g_n^* - \bar{g}_n^*||_{[a,b]} + ||\bar{g}_n - E \bar{g}_n||_{[a,b]} + ||E \bar{g}_n - E g_n^*||_{[a,b]} \).

Since \( E|Y|^p < \infty \) implying \( E|e|^p < \infty \), then there exists a full set \( \Omega_0 \) such that for each \( \omega \in \Omega_0 \) there exists a finite positive integer \( N_\omega \) and for all \( n \geq N_\omega \) we have

\[ e_n(\omega) = \bar{e}_n(\omega). \]

Hence for all \( n \geq N_\omega \),

(2.10) \( |g_n(x) - \bar{g}_n(x)| \leq \sum_{i=1}^{N_\omega} |e_i - \bar{e}_i| \int_{x_{i-1}}^{x_i} a_n^{-1} k\left(\frac{x-z}{a_n}\right) dz \)

\[ \xrightarrow{\text{w.p.1}} C(N_\omega) \delta \cdot n^{-1} ||k||_{[-L,L]}, \]
where $0 < C(N_{\omega}) < \infty$. Therefore, in connection with condition (iv), we have

\begin{equation}
(n^{\gamma}/\beta_n \log n) \|g_n^* - \overline{g}_n\|_{[a,b]} \xrightarrow{w.p.1} 0, \text{ as } n \to \infty.
\end{equation}

Finally,

\begin{align*}
|E_{g_n}(x) - E_{g_n}^*(x)| &= \sum_{i=1}^{n} |e_i| I(|e_i| > i^{1/p}) \int_{x_{i-1}}^{x_i} a_n^{-1} k\left(\frac{x - z}{a_n}\right) dz dP(e_i) \\
&\leq \sum_{i=1}^{n} i^{-(p-1)/p} \int |e_i| dP(e_i) \int_{x_{i-1}}^{x_i} a_n^{-1} k\left(\frac{x - z}{a_n}\right) dz \\
&\leq C(\sum_{i=1}^{n} i^{-(p-1)/p}) \delta_n a_n^{-1} \|k\|_{[-L,L]}, \quad 0 < C < \infty.
\end{align*}

$a C' n^{1/p} \delta_n a_n^{-1}$, for large $n$ and $0 < C' < \infty$.

Consequently,

\begin{equation}
(n^{\gamma}/\beta_n \log n) \|E_{g_n} - E_{g_n}^*\|_{[a,b]} \to 0
\end{equation}

as $n \to \infty$ by condition (iv). Now, upon the substitution of (2.6), (2.11), and (2.12) into (2.9), the desired result (2.1) is established. \(\square\)

With different combination of $a$, $\beta$, $\gamma$, and $p$, various interesting rates for strong uniform convergence may be achieved. For example, take $\alpha = \beta = 1$, $a_n = n^{-\gamma}$, $\gamma < 1/2$, and $\delta_n = O(n^{-1})$. If $E|e|^2 < \infty$, then

\begin{align*}
\|g_{1n} - g\|_{[a,b]} \xrightarrow{w.p.1} 0, \text{ as } n \to \infty.
\end{align*}

As another example, take $\alpha = \beta = 1$, $a_n = n^{-1/3}$, and $c/n \leq \delta_n \leq c'/n$ for $cc' \neq 0$. If $E|e|^3 < \infty$, then

\begin{align*}
(n^{1/3}/\beta_n \log n) \|g_{1n} - g\|_{[a,b]} \xrightarrow{w.p.1} 0, \text{ as } n \to \infty.
\end{align*}
3. Isotonic Estimation of $g(x)$. In the case that the regression
function $g(x)$ is known to be nondecreasing (or nonincreasing) a question
which arises naturally is how to reasonably modify the estimators
$g_{1n}(x)$, $i = 1, 2, 3$, so that the new estimators will preserve the monotonicity
and also enjoy some optimal properties. In this section, we will only
consider a modification of $g_{1n}(x)$ using the method of Barlow, et al (1972)
when $g$ is nondecreasing; the same technique applies to those of $g_{2n}(x)$ and
$g_{3n}(x)$. The case of nonincreasing $g$ can be treated by symmetry.

**Definition.** Let $S$ be a finite set $\{t_1, \ldots, t_n\}$ with the simple
ordering $t_1 < t_2 < \ldots < t_n$. A real-valued function $f$ on $S$ is said to be
isotonic if $x, y \in S$ and $x < y$ imply $f(x) \leq f(y)$. Let $g(x)$ and $w(x) > 0$ be
given functions on $S$. Then an isotonic function $g^{**}$ is said to be an
isotonic regression of $g$ with weight $w$, relative to the simple ordering
$t_1 < \ldots < t_n$, if it minimizes, in the class of isotonic functions $f$ on $S$,
the sum

$$(3.1) \quad \sum_{x \in S} [g(x) - f(x)]^2 w(x).$$

Now, for each $n$, consider the set $S_n = \{t_{ni}\} \subset [a, b]$ as a grid on $[a, b]$
such that $S_n$ is dense in $[a, b]$ as $n \to \infty$. Let each $t_{ni}$ be assigned a positive
weight $w(t_{ni})$. Then a max-min type isotonic regression of $g$ with weight $w$
on the set $S_n$ is given by

$$(3.2) \quad g_{n}^{**}(t_{ni}) = \max \min \frac{t}{\sum_{s \leq i} \sum_{t \geq j} g_{1n}(t_{nj})w(t_{nj})/\sum_{j=s}^t w(t_{nj})}. $$

Furthermore, for all $x \in [x_{nj}, x_{n\cdot j+1}]$, we define

$$(3.3) \quad g_{n}^{**}(x) = g_{n}^{**}(t_{nj}). $$
Consequently, $g_n^{**}$ is a nondecreasing step function on $[a, b]$. Therefore, it follows from Theorem 1.6 of Barlow, et al (1972) and Theorem 4 that, for each $x \in [a, b]$,

$$g_n^{**}(x) \xrightarrow[w.p.]{} g(x) \text{ as } n \to \infty.$$ 

As an example, we may take

$$S_n = \{ t_{nj} = a + (b - a)j2^{-n}, \ j = 0, 1, \ldots, 2^n \}$$

as a grid on $[a, b]$. 
REFERENCES


Consider the regression model $Y_i = g(X_i) + e_i$, $i = 1, ..., n$, where $g$ is an unknown function defined on the interval $[0, 1]$. Fix $0 < x_1 \leq x_2 \leq ... \leq x_n \leq 1$, with $\max(x_i - x_{i-1}) = O(n^{-1})$, and assume $\{e_i\}$ are iid random variables with zero mean and finite variance. Priestley and Chao (J. Roy. Statist. Soc. B 34 (1972), 385-392) and Clark (J. Roy. Statist. Soc. B 39 (1977), 107-113) proposed estimators $g_{2n}$ and $g_{3n}$.
respectively, for $g$. In this paper, asymptotic behaviors of $g_{2n}$ and $g_{3n}$ are studied utilizing the properties of the new estimator $g_{1n}(x) = \sum_{i=1}^{n} Y_i \{ K[(x - x_{i-1})/a_n] - K[(x - x_i)/a_n] \}$ where $K(u)$ is a known absolutely continuous cdf with pdf $k(u)$. It is shown that $g_{1n}$, $g_{2n}$, and $g_{3n}$ are asymptotically equivalent in various senses. Moreover, consistent results are established and rates of uniform convergence obtained. For example, if $E|Y_i|^3 < \infty$, if both $g$ and $k$ are Lipschitz of order 1, and if $\{g_n\}$ is any diverging sequence as $n \to \infty$, then, for $0 < a \leq b \leq 1$,

$$\frac{n^{1/3}}{b \log n} \sup_{a \leq x \leq b} |g_{1n}(x) - g(x)|_{w,p,1} \to 0, \text{ as } n \to \infty.$$  

Finally, when $g$ is monotone, a strong consistent isotonic estimator $g^*_n$ is considered.