EXPONENTIAL BOUNDS OF MEAN ERROR FOR THE NEAREST NEIGHBOR ESTIMATES OF REGRESSION FUNCTIONS

L. C. Zhao

Center for Multivariate Analysis
University of Pittsburgh

Approved for public release; distribution unlimited.
EXPONENTIAL BOUNDS OF MEAN ERROR FOR THE
NEAREST NEIGHBOR ESTIMATES OF
REGRESSION FUNCTIONS

L. C. Zhao

November 1985

Technical Report 85-44

Center for Multivariate Analysis
Fifth floor, Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

*Research sponsored by the Air Force Office of Scientific Research (AFSC) under contract #F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.
EXPONENTIAL BOUNDS OF MEAN ERROR
FOR THE NEAREST NEIGHBOR ESTIMATES
OF REGRESSION FUNCTIONS

L. C. Zhao

ABSTRACT

Let \((X, Y), (X_1, Y_1), \ldots, (X_n, Y_n)\) be i.i.d. \(\mathbb{R}^r \times \mathbb{R}\)-valued random vectors with \(E|Y|<\infty\), and let \(m_n(x)\) be a nearest neighbor estimate of the regression function \(m(x) = E(Y|X=x)\). In this paper, we establish an exponential bound of the mean deviation between \(m_n(x)\) and \(m(x)\) given the training sample \(Z^n = (X_1, Y_1, \ldots, X_n, Y_n)\), under the conditions as weak as possible. This is a substantial improvement on Beck's result.

Key words. Regression function, nearest neighbor estimate, exponential bound, mean error, training sample.
1. INTRODUCTION

Let \((X,Y), (X_1,Y_1), \ldots, (X_n,Y_n)\) be i.i.d. \(\mathbb{R}^d \times \mathbb{R}\)-valued random vectors with \(E|Y|<\infty\). To estimate \(m(x) = E(Y|X=x)\), the regression function of \(Y\) with respect to \(X\), Stone (1977) and others proposed the so-called weight estimation

\[
m_n(x) = \sum_{j=1}^{n} W_{nj}(x)Y_j,
\]

where \(W_{nj}(x) = W_{nj}(x_1, \ldots, x_n)\) is a Borel-measurable function of its arguments. Let \(V_{nj}, j = 1, \ldots, n\), be non-negative real number such that \(\sum_{j=1}^{n} V_{nj} = 1\). For suitable-chosen metric \(\|a-b\|\) on \(\mathbb{R}^d\) (such as \(L^2\) or \(L^\infty\)), rearrange \(X_j, j = 1, \ldots, n:\)

\[
\|X_1^x - x\| \leq \|X_2^x - x\| \leq \ldots \leq \|X_n^x - x\|
\]

(ties are broken by comparing indices), and set

\[
m_n(x) = \sum_{j=1}^{n} V_{nj}Y_j^x.
\]

Then we obtain the nearest neighbor (NN) estimates of \(m(x)\).

Many scholars studied convergence problem of these estimates from different points of view. (For the universal consistency, one can refer to, for example, Stone (1977). For the pointwise moment-consistency, see Devroye (1981). For the pointwise a.s. consistency, see Devroye (1981), Zhao and Bai (1984)). In this paper, we study another convergency of these estimates.

Write \(X^n = (X_1, \ldots, X_n), Y^n = (Y_1, \ldots, Y_n)\) and \(Z^n = (X^n, Y^n)\). Let \(g_n = g_n(x, Z^n)\) be an estimate of \(m(x)\). In some problems, we are interested in the following mean deviation of \(g_n\) given the training sample \(Z^n:\)

\[
D(g_n) = E\{|g_n(x, Z^n) - m(x)| | Z^n\} = \int_{\mathbb{R}^d} |g_n(x, Z^n) - m(x)| Q(dx),
\]

where \(Q\) denotes the distribution of \(X\).
Take \( k = k_n \leq n \), and put

\[
\hat{m}_n(x) = \frac{1}{k} \sum_{j=1}^{k} y_{x_j}.
\]

For this class of estimates, Beck (1979) established the following theorem:

Suppose that the following conditions are satisfied:

\[
(i) \ Y \text{ is bounded.}
\]
\[
(ii) \ m(x) \text{ is continuous on } \mathbb{R}^d.
\]
\[
(iii) \ Q \text{ has a continuous density } f.
\]
\[
(iv) \ k \to \infty \text{ and } k/n \to 0 \text{ as } n \to \infty.
\]

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

\[
P\{D(\hat{m}_n) > \varepsilon\} \leq e^{-cn}
\]

where \( C > 0 \) is a constant independent of \( n \).

This theorem deals only with a special case of NN estimates, and the assumptions are rather restrictive. Recently, we substantially improved this result. We established the following:

Theorem 1. Let \( m_n(x) \) be a NN estimate of \( m(x) \) defined by (2) and (3).

Suppose that the following conditions are satisfied:

\[
(i) \ Y \text{ is bounded.}
\]
\[
(ii) \ Q \text{ has a density } f.
\]
\[
(iii) \ \text{There exists a sequence of integers } k = k_n \text{ such that}
\]
\[
k \to \infty, \ k/n \to 0,
\]
\[
\sup_n \{k_{\max} \leq j < k_{n+1} \} < \infty \text{ and } \sum_{j=k+1}^{n} V_{n_j} \to 0.
\]

Then for any given \( \varepsilon > 0 \), we have

\[
P\{D(m_n) > \varepsilon\} \leq e^{-cn},
\]

where \( C > 0 \) is a constant independent of \( n \).
Note that the special case considered by Beck is included in this theorem. Besides, this theorem gives a substantial improvement of Beck's result, by getting rid of the continuity requirement of \( m(x) \) and \( f(x) \), the density of \( Q \).

2. SOME LEMMAS.

Theorem 1 is valid for the \( L_2 \) norm or \( L_\infty \) norm on \( \mathbb{R}^d \), here we only give the proof for \( L_\infty \) norm. For simplicity, we make the following convention: \( \varepsilon_1, \varepsilon_2, ..., \alpha, \beta_1, \beta_2, \delta, \) etc., are all constants independent of \( n \). \( I_A \) or \( I(\mathbb{A}) \) denotes the indicator of a set \( A \). \( \#(A) \) denotes the cardinal of set \( A \). \( S_{x,\rho} = \{ u \in \mathbb{R}^d : ||u-x|| \leq \rho \} \).

\( Q \) and \( \lambda \) denote the outer measure generated by \( Q \) and the Lebesgue measure \( \lambda \) (on \( \mathbb{R}^d \)), respectively. We need the following lemmas in the sequel.

Lemma 1 (Besicovitch Covering Lemma). Let \( E \) be bounded subset of \( \mathbb{R}^d \), and let \( K \) be a family of cubes covering \( E \) which contains a cube \( D_x \) with center \( x \) for each \( x \in E \). Then there exist points \( \{x_k\} \) in \( E \) such that

(i) \( E \subseteq \bigcup_{k} D_{x_k} \).

(ii) there exists a constant \( \sigma \) depending only on \( d \) such that \( \sum_{k} I(D_{x_k}) \leq \sigma \).

Refer to Wheeden and Zygmund (1977), pp. 185-187.

Let \( Q_n \) be the empirical measure of \( X_1, ..., X_n \), and \( T > 0 \) be a given constant. Fix \( \delta \in (0, 1/2\sigma) \) and assume that \( h = h_n \in (0, 1) \). Set

\[
G_n^* = \{ x \in S_{O, T} : Q_n(S_x, h) < \delta Q(S_x, h) \},
\]

and

\[
E^* = \{ x \in S_{O, T} : \beta_1(2\rho)^d < Q(S_x, h) < \beta_2(2\rho)^d \}
\]

for any \( \rho \in (0, 1) \),

where \( \beta_1 > 0 \) and \( \beta_2 > 0 \) are constants to be chosen later.
LEMMA 2. Suppose that $Q$ has a density $f$. Then for any $\varepsilon > 0$, we can choose $\beta_1$ small enough and $\beta_2$ large enough such that $Q^*(S_0, T-E^*) < \varepsilon$.

Note that for any Borel-measurable set $E \subseteq E^*$, we have

$$\beta_1 \leq f(x) \leq \beta_2, \text{ for almost all } x \in E^*(\lambda).$$

LEMMA 3. Suppose that $Q$ has a density $f$, $h = h_n \in (0,1)$ and $nh^d \to \infty$. Then for any given $\varepsilon > 0$, we have

$$P(Q^*(G^*_n) > \varepsilon) < e^{-C_1 n}.$$ 

Lemmas 2 and 3 can be deduced from Lemma 1. For the proof, see Zhao (1985).

Lemma 4. Suppose that $\int_R |g(x)|Pf(dx) < \infty$ for some $p > 0$, then

$$\lim_{h \to 0} \int S_{x,h} \frac{|g(u)-g(x)|Pf(du)/f(S_{x,h})}{0} = 0$$

for almost all $x(F)$.

Refer to Wheeden and Zygmund (1977), p. 191, example 20.

3. Proof of Theorem 1

Suppose that $|Y| \leq M$. Then

$$\int \left| \sum_{j \geq k} V_n (Y_j - m(x)) \right| Q(dx) \leq 2M \sum_{j \geq k} V_n j \to 0$$

as $n \to \infty$. Without loss of generality, we can assume $\sum_{j \geq k} V_n j = 0$ for any $n$. It is enough to prove that for each fixed $T > 0$,

$$P(\int_{S_0, T/2} |m_n(x)-m(x)|Q(dx) \geq \varepsilon) < e^{-C_1 n}.$$ 

By Lemma 2, there exists $\beta_1 = \beta_i(\varepsilon)$, $i = 1, 2$, and a compact set $E \subseteq E^*$ such that
\[(11) \quad Q(S_{0,T-E}) < \epsilon/8M, \]

where \(E^*\) is defined by (9).

Fix \(\delta \in (0, \frac{1}{2})\), and take \(\alpha \geq (2^{d-1})^{-1}\). Set

\[h = h_n = (\alpha k/n)^{1/d},\]

then \(h \to 0\) and \(nh^d \to \infty\) as \(n \to \infty\).

By Lemma 3, there exists a compact set \(H_n\) such that with \(h\) as above

\[(12) \quad H_n \subset \{x \in S_{0,T}: Q_n(S_x,h) \geq \delta Q(S_x,h)\}\]

and

\[(13) \quad P\{Q(S_{0,T-H_n}) > \epsilon/8M\} < \frac{1}{n}.\]

For \(x \in H_n \cap E\), \(Q_n(S_{x,h}) \geq \delta Q(S_x,h) \geq \delta^2 \delta \lambda(S_x,h) = \delta^2 \delta \delta k/n \geq k/n\), so that \(X_1^x, X_2^x, \ldots, X_k^x\) all fall into \(S_{x,h}\).

Partition \(R^d\) into sets with the form \(\prod (i_j-1)h, i_j h), \) where \(i_1, \ldots, i_d = 0, \pm 1, \ldots\). Call the partition \(\psi\). Set \(\psi' = \{B \in \psi, B \subset S_{0,T}\}\). For \(B \in \psi'\), put

\[\tilde{W}(B) = \{B' \in \psi', \rho(B,B') < 3h\}, \quad W(B) = \bigcup_{B' \in \tilde{W}(B)} B'.\]

where \(\rho(B,B') = \inf \{|x-x'|: x \in B, x' \in B'\}\). Then there exists a constant \(C_d\) such that for any \(B \in \psi'\) we have \#(\tilde{W}(B)) \leq C_d\). It is easy to show by induction that, \(\psi'\) can be divided into \(C_2(\leq C_d^2)\) disjoint subsets \(\psi_i, i=1, \ldots, C_2\), such that for any two sets \(B_1, B_2\) in the same \(\psi_i\), we have

\[W(B_1) \cap W(B_2) = \emptyset.\]

Denote by \(B(x)\) the cube \(B \in \psi\) which contains \(x\). If \(x \in H_n \cap E\) and \(B(x) \in \psi'\), then for any \(u \in B(x)\), we have \(S_{x,h} \subset S_u, 2h \subset W(B(x))\), so that, from \(Q_n(S_{x,h}) \geq k/n\) it follows that \(X_1^u, \ldots, X_k^u\) are also contained in \(W(B(x))\). If we write
then, as mentioned above, for any \( B \in A_n \), \( W(B) \) contains the \( k \) nearest neighbors of each \( x \in B \). Further, we set \( H_i = A_n \cap \psi_i \), \( i = 1, 2, \ldots, C_2 \). It is easy to see that

\[
\int_{S_0, T/2} |m_n(x) - m(x)| Q(dx) \leq \int_{S_0, T-E} + \int_{S_0, T-H_n} + \int_{H_n \cap E \cap S_0, T/2}.
\]

By (11), we have

\[
\int_{S_0, T-E} |m_n(x) - m(x)| Q(dx) \leq 2MQ(S_0, T-E) < \varepsilon/4.
\]

By (13),

\[
P(\int_{S_0, T-H_n} |m_n(x) - m(x)| Q(dx) \geq \varepsilon/4) \\
\leq P(Q(S_0, T-H_n) \geq \varepsilon/8M) < e^{-c_1 n}.
\]

Hence to prove (10), it is enough to prove that

(14) \[ P(\int_{H_n \cap E \cap S_0, T/2} |m_n(x) - m(x)| Q(dx) \geq \varepsilon/2) < e^{-C_3 n} \]

For large \( n \),

\[
\int_{H_n \cap E \cap S_0, T/2} |m_n(x) - m(x)| Q(dx) \\
\leq \sum_{B \in A_n} \int_{B \cap E} |m_n(x) - m(x)| Q(dx) \\
\leq \sum_{i=1}^{C_2} \sum_{B \in H_i} \int_{B \cap E} |m_n(x) - m(x)| Q(dx).
\]

Put

\[
\bar{m}_n(x) = \frac{1}{k} \sum_{j=1}^{k} m(X_j),
\]

\[
I_{n1} = \sum_{B \in H_i} \int_{B \cap E} |m_n(x) - \bar{m}_n(x)| Q(dx).
\]
\[ J_{ni} = \sum_{B^i \in H_1} \int_{B \cap E} \bar{m}_i(x) \, dx, \quad i = 1, \ldots, C. \]

\[ \phi(B) = \int_{B \cap E} \frac{1}{k} \sum_{j=1}^{k} V_n j(y_j - m(x_j)) \, dx, \]

\[ d_{ni} = \#\{B^i \in H_1, \phi(B) > \varepsilon/(3C^2)\}, \quad i = 1, \ldots, C. \]

To prove (14), it is enough to show that, for each \( i, 1 \leq i \leq C \), we have

\[ P[I_{ni} \geq \varepsilon/(4C^2)] < e^{-C_4 n}. \]

\[ P[J_{ni} \geq \varepsilon/(4C^2)] < e^{-C_5 n}. \]

For almost all \( x \in B \cap E(\lambda) \), \( f(x) \leq \beta_2 \). Hence,

\[ I_{ni} \leq \varepsilon/(8C^2) + 2M d_{ni} \beta_2^2 \alpha k/n. \]

Write \( C_6 = e(16MC^2 \beta_2)^{-1} \), then

\[ P[I_{ni} \geq \varepsilon/(4C^2)] \leq P[d_{ni} \geq C_6 n/k]. \]

Now we proceed to prove that, for any \( B \in H_1 \),

\[ P[\phi(B) \geq \varepsilon/8C^2 | X^n] < e^{-C_7 k}, \]

where \( X^n = (X_1, \ldots, X_n) \) is defined as before.

For any \( \varepsilon_1 > 0 \) and \( s > 0 \), by Jensen's inequality we have

\[ P[\phi(B) \geq \varepsilon_1 | X^n] \leq e^{-s \varepsilon_1} \mathbb{E} \{ \exp(s \phi(B)) | X^n \} \]

\[ \leq e^{-s \varepsilon_1} \int_{B \cap E} \mathbb{E} \{ \exp(s \sum_{j=1}^{k} V_n j(y_j - m(x_j))) | X^n \} \, dx / Q(B \cap E). \]
When \( \{X_j^X, j \leq k\} \) is given, \( Y_1^X, \ldots, Y_k^X \) are independent. From this and the inequality \( |e^{t} - 1 - t| \leq \frac{1}{2} t^2 |e^t| \) for any real \( t \), it follows that,

\[
\begin{align*}
E\{\exp(\sum_{j=1}^{k} V_{nj} [Y_j^X - m(X_j^X)]) | X^n \} \\
&= \prod_{j=1}^{k} E\{\exp(\sum_{j=1}^{k} V_{nj} [Y_j^X - m(X_j^X)]) | X_j^X \} \\
&\leq \prod_{j=1}^{k} \{1 + s^2 C^2 k^{-1} \exp(2sCk^{-1})\} \\
&\leq \exp\{s^2 C^2 k^{-1} \exp(2sCk^{-1})\}.
\end{align*}
\]

Here we have written \( C^9 = \sup_{n} \{k \max_{j \leq k} V_{nj}\} \) and \( C^8 = C^9 M \). In the same way,

\[
E\{\exp(\sum_{j=1}^{k} V_{nj} [m(X_j^X) - Y_j^X)]) | X^n \} \\
\leq \exp\{s^2 C^2 k^{-1} \exp(2sCk^{-1})\}.
\]

In view of (20), we get

\[
P(\phi(B) \geq \varepsilon_1 | X^n) \leq 2 \exp\{-s\varepsilon_1 + s^2 C^2 k^{-1} \exp(2sCk^{-1})\}
\]

Take \( s = \mu k \) with \( \mu \) being small enough, we have

\[
P(\phi(B) \geq \varepsilon_1 | X^n) < e^{-C_{10} k}.
\]

This is just (19).

Since for each \( B \in H^i_1 \), \( W(B) \) contains the \( k \) nearest neighbors of each \( x \in B \), and \( W(B_1) \cap W(B_2) = \emptyset \) for any \( B_1, B_2 \in H^i_1 \), we see that when \( X^n = (X_1^x, \ldots, X_n^x) \) is given, \( \{\phi(B), B \in H^i_1\} \) is a group of conditionally independent variables. Put \( G(B) = \{\phi(B) \geq \varepsilon_1\} \). Then by (19) and \( #(H^i_1) \leq #(\psi') \leq C_{11} n/k \), we have
From (18) and (21) it follows (16) is valid.

Now we proceed to prove (17). As mentioned above, for each $B \in \mathcal{H}_1$, $X_1, \ldots, X_k$ all fall into $W(B)$. Noticing the conditions imposed on $\nu_j$'s, we see that

$$\sum_{B \in \mathcal{H}_1} \nu_j = \sum_{B \in \mathcal{H}_1} \mathbb{E}(\mathbb{E}_{X_j}(m(X_j)-m(x))|Q(dx))$$

$$\leq Cg^{-1}\sum_{B \in \mathcal{H}_1} \sum_{j=1}^{n} I_W(B)(X_j) \mathbb{E}(\mathbb{E}_{X_j}(m(X_j)-m(x))|Q(dx))$$

$$= Cg^{-1}\sum_{B \in \mathcal{H}_1} \sum_{j=1}^{n} I_W(B)(X_j)Z_B(X_j),$$

where

$$Z_B(u) = \mathbb{E}(\mathbb{E}_{X_j}(m(u)-m(x))|Q(dx)) \leq 2M e^{2} \leq e^{2}.$$

Here, the following facts are used: $|m(x)| \leq M$, $f(x) \leq \beta_2$ for $x \in B \cap E$ and, $\lambda(B) \leq h^d = ak/n$.

Put $\varepsilon_2 = Cg^{-1}$. To prove (17), it suffices to prove that

$$P(\sum_{B \in \mathcal{H}_1} \sum_{j=1}^{n} I_W(B)(X_j)Z_B(X_j) \geq 2k\varepsilon_2^2) \leq e^{-C_1n}.$$
Let $N$ be a Poisson random variable with parameter $n$, which is independent of $X_1, X_2, \ldots$. If $|N-n| < n \varepsilon_3 = n\varepsilon_2/(2M_2\alpha)$, then by (23)

$$
|\sum_{B \in \Psi} \left( \sum_{j=1}^{n} I_W(B)(X_j)Z_B(X_j) - \sum_{j=1}^{n} I_W(B)(X_j)Z_B(X_j) \right)|
\leq |N-n|2M_2\alpha k/n < \varepsilon_2^k.
$$

It follows that

$$(25) \quad P\left\{ \sum_{B \in \Psi} \sum_{j=1}^{n} I_W(B)(X_j)Z_B(X_j) > 2k\varepsilon_2 \right\}
\leq P\left\{ |N-n| > n \varepsilon_3 \right\} + P\left\{ \sum_{B \in \Psi} \sum_{j=1}^{N} I_W(B)(X_j)Z_B(X_j) > k\varepsilon_2 \right\}$$

It is easy to show that

$$(26) \quad P\left\{ |N-n| > n \varepsilon_3 \right\} < e^{-C_14n}.$$ 

Since $W(B)$, $B \in \Psi$, are disjoint, we see that for $t > 0$,

$$(27) \quad P\left\{ \sum_{B \in \Psi} \sum_{j=1}^{n} I_W(B)(X_j)Z_B(X_j) > k\varepsilon_2 \right\}
\leq e^{-t\varepsilon_2^k} \sum_{\ell=0}^{\infty} \frac{e^{-n \ell}}{\ell!} \left( E\{ \exp(t\sum_{B \in \Psi} I_W(B)(X_1)Z_B(X_1)) \} \right)^\ell
\leq e^{-t\varepsilon_2^k} \sum_{\ell=0}^{\infty} \frac{e^{-n \ell}}{\ell!} \left( \sum_{B \in \Psi} \left[ W(B) e^{tZ_B(u)} Q(du) + 1 - Q(U \in \Psi) \right] \right)^\ell
\leq \exp\{-t\varepsilon_2^k + n\sum_{B \in \Psi} \int W(B) e^{tZ_B(u)} Q(du) - 1\} Q(du)$$

Now we proceed to show that

$$(28) \quad \limsup_{n \to \infty} \sum_{B \in \Psi} \int W(B) \left[ \exp\left( \frac{\etaZ_B(u)}{k^2} \right) - 1 \right] Q(du) = 0.$$ 

By (23), there exist constants $C_{15}, C_{16}$ such that
\[ \frac{n}{k} z_B(u) \leq c_{15} \]

and

\[ \exp\left(\frac{n}{k} z_B(u)\right) - 1 \leq c_{16} \frac{n}{k} z_B(u). \]

To prove (28), it suffices to show that

(29) \[ \limsup_{n \to \infty} \frac{1}{k} \sum_{B \in \mathcal{V}_k} \int_{B \cap E} Q(dx) \int_{W(B)} \left| m(u) - m(x) \right| Q(du) = 0. \]

Assume that \( B \in \mathcal{V}_k \), \( B \cap E \neq \emptyset \) and \( x \in B \cap E \), then \( W(B) \subseteq S_{x, 5h} \), where \( h = (\alpha k/n)^{1/d} \). By Lemma 2,

\[ Q(S_{x, 5h}) \leq B_2(10h)^d = 10^d B_2 \alpha k/n. \]

Put \( C_{17} = 10^d B_2 \alpha \), then

(30) \[ \frac{1}{k} \sum_{B \in \mathcal{V}_k} \int_{B \cap E} Q(dx) \int_{W(B)} \left| m(u) - m(x) \right| Q(du) \leq C_{17} \sum_{B \in \mathcal{V}_k} \int_{B \cap E} Q(dx) \int_{S_{x, 5h}} \left| m(u) - m(x) \right| Q(du)/Q(S_{x, 5h}) \]

By Lemma 4, for almost all \( x(\Omega) \),

\[ \lim_{n \to \infty} \int_{S_{x, 5h}} \left| m(u) - m(x) \right| Q(du)/Q(S_{x, 5h}) = 0. \]

Further, for \( x \in S(\Omega) \), the support of \( Q \), we have

\[ \int_{S_{x, 5h}} \left| m(u) - m(x) \right| Q(du)/Q(S_{x, 5h}) \leq 2M \]

Hence, by the dominated convergence theorem, (29) is valid. Thus (28) is proved.

Take \( t = n/k \) in (27), we have
From (25), (26) and (31), it follows that (24) holds, and (17) is valid. From (16) and (17), Theorem 1 is proved.
REFERENCES


Exponential bounds of mean error for the nearest neighbor estimates of regression functions

L. C. Zhao

Center for Multivariate Analysis
515 Thackeray Hall
University of Pittsburgh, Pittsburgh, PA 15260

Air Force Office of Scientific Research
Boiling Air Force Base, DC 20332

Approved for public release; distribution unlimited

Let \((X,Y)(X_1,Y_1),\ldots,(X_n,Y_n)\) be i.i.d. \(R^n \times R\)-valued random vectors with \(E|Y|<\infty\), and let \(m_n(x)\) be a nearest neighbor estimate of the regression function \(m(x) = E(Y|X=x)\). In this paper, we establish an exponential bound of the mean deviation between \(m_n(x)\) and \(m(x)\) given the training sample \(Z^n = (X_1,Y_1),\ldots,(X_n,Y_n)\), under the conditions as weak as possible. This is a substantial improvement on Beck's result.
END

DTIC

5-86