第3卷 第1期　　数学研究与评论　　Vol.3 No.1
1983年 1月　　JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION　　Jan. 1983

Uniform Convergence Rates of the Nearest Neighbor Density Estimates*

H. R. Chen (Chen Xiru 陈希孺)

Let $X_1, \cdots, X_n$ be iid. samples drawn from a population with probability density function f and distribution function F. There are a lot of discussions concerning the problem of estimating f from these samples. In 1965, Loftsgarden and Quesenberry[1] proposed the following scheme: Choose a positive integer $k=k_n$ depending upon n such that $1\leqslant k_n\leqslant n$. Find the smallest number $a_n(x)=a_n(x; X_1, \cdots, X_n)$ satisfying the condition

$$\#(\{i: 1\leqslant i\leqslant n;\ x-a_n(x)\leqslant X_i<x+a_n(x)\})\geqslant k_n$$

where $\#(A)$ denotes the number of elements contained in the set A. Define

$$\hat{f}_n(x)=k_n/(2na_n(x)) \tag{1}$$

as the estimate of $f(x)$.

A number of authors have studied the consistency of this estimate—sometimes known as the Nearest Neighbor Estimate.

The best result was obtained in 1977 by Devroye and Wagner, who showed in [2] that under the conditions

a. f is uniformly continuous on R^m,

b. $\lim\limits_{n\to\infty} k_n/n=0$, $\lim\limits_{n\to\infty} \log n/k_n=0$,

then as $n\to\infty$ with probability one, we have

$$\sup_x|\hat{f}_n(x)-f(x)|\to 0. \tag{2}$$

From this result, the convergence rate of (2) naturally presents itself. This problem is of much interest, for one thing, a similar problem for the classical kernel estimate has been studied extensively in the literature. In [3], the author has obtained some results in this respect:

1. No convergence rate of (2) can be established without some further restrictions imposed on f, beyond that of being uniformly continuous.

2. In case $m=1$, supposing that f satisfies Lipshitz condition, for some properly chosen k_n we can get

* Received Nov. 24, 1981.

$$\sup_x |\hat{f}(x) - f(x)| = O(n^{-1/6}(\log\log n)^{1/6}), \quad a.s. \tag{3}$$

3. Also in case $m=1$, for arbitrarily chosen k_n, one can find a density function f satisfying Lipshitz condition, yet the assertion

$$\sup_x |\hat{f}(x) - f(x)| = O(n^{-1/4}(\log\log n)^{-1/4}) \ a.s.$$

is not true.

Based on these results, the present author advanced a conjecture that for f satisfying Lipshitz condition, the rate presented in (3) can be improved to $O(n^{-1/4+\varepsilon})$ for each $\varepsilon>0$. This means that under the above condition, the exponential 1/4 is best and can no longer be improved. The purpose of this paper is to prove this result and its extension to probability densities satisfying the δ-th order Lipshitz condition $(0<\delta\leqslant 1)$.

The main result of this paper can be formulated as follows:

Theorem Let $0<\delta\leqslant 1$ and $\mathscr{F}_\delta$ denote the the family of all probability density function in R^1 satisfying the Lipshitz condition of δ-th order. If we choose

$$k = k_n = [n^{2\delta/(1+3\delta)}] \tag{4}$$

and define $\hat{f}_n(x)$ by (1), then for any $c_n\to\infty$ we have

$$\sup_x |\hat{f}_n(x) - f(x)| = O(n^{-\delta/(1+3\delta)}(\log n)^{1/2}c_n), \quad a.s. \tag{5}$$

for any $f\in\mathscr{F}_\delta$. On the other hand, for any $\delta\in(0,1]$, one can find $f\in\mathscr{F}_\delta$ such that for any choice of k_n, the assertion

$$\sup_x |\hat{f}_n(x) - f(x)| = O(n^{-\delta/(1+3\delta)}) \quad a.s.$$

is not true.

Proof The proof makes use of a special case of a powerful inequality given by Devroye and Wagner in [4], which we formulate below as a lemma.

Lemma. 1 Suppose that $X_1,\cdots,X_n$ are independent one-dimensional random variables with a common distribution function F. Denote by F_n the empirical distribution function of $X_1,\cdots,X_n$. Then, for any $\varepsilon>0$, $0\leqslant B\leqslant 1/4$ and $n\geqslant\max(B^{-1}, 8B\varepsilon^{-2})$, we have

$$P(\sup\{|\overline{F_n(b)-F_n(a)} - \overline{F(b)-F(a)}| : 0\leqslant F(b)-F(a)\leqslant B\}\geqslant\varepsilon)$$
$$\leqslant 16n^2\exp\left(\frac{-n\varepsilon^2}{64B+4\varepsilon}\right) + 8n\exp\left(-\frac{nB}{10}\right), \tag{6}$$

Turning to the proof of the theorem, we choose

$$B = 2n^{-(1+\delta)/(1+3\delta)}$$

and use $n^{-(1+2\delta)/(1+3\delta)}(\log n)^{1/2}c_n$ to replace ε in (6) ($\varepsilon>0$ given). Note that $B\leqslant 1/4$, $n>B^{-1}$ and

$$8B\varepsilon^{-2} = 16n/[(\log n)c_n^2\varepsilon^2] < n$$

all for n large. We get for these n

$$P(n^{(1+2\delta)/(1+3\delta)}(\log n)^{-1/2}c_n^{-1}\sup_x\{|\overline{F_n(b)-F_n(a)}-\overline{F(b)-F(a)}|:$$
$$0\leqslant F(b)-F(a)\leqslant 2^{-(1+\delta)/(1+3\delta)}\}\geqslant\varepsilon)$$
$$\leqslant 16n^2\exp\left(-\frac{c_n^2\varepsilon^2\log n\cdot n^{(1+\delta)/(1+3\delta)}}{128n^{-(1+\delta)/(1+3\delta)}+4\varepsilon c_n(\log n)^{1/2}n^{-(1+2\delta)/(1+3\delta)}}\right)$$
$$+8n\exp(-5^{-1}n^{2\delta/(1+3\delta)}) \tag{7}$$

by employing the lemma. As $\delta>0$, $\varepsilon>0$ are fixed and $c_n\to\infty$, one sees that the right-hand side of (7) is of the order $O(n^{-2})$. Hence the series with a general term as the left-hand side of (7), summed up from $n=n_0$ (n_0 sufficiently large) to ∞, is convergent. By the arbitrariness of $\varepsilon>0$, it follows that

$$\sup_x\{|\overline{F_n(b)-F_n(a)}-\overline{F(b)-F(a)}|:0\leqslant F(b)-F(a)\leqslant 2n^{-(1+\delta)/(1+3\delta)}\}$$
$$=o(c_n(\log n)^{1/2}n^{-(1+2\delta)/(1+3\delta)}).\quad a.s. \tag{8}$$

Since $\delta>0$, we have*)

$$c_n(\log n)^{1/2}n^{-(1+2\delta)/(1+3\delta)}=o(n^{-(1+\delta)/(1+3\delta)}).$$

From (8), one sees that with probability one, in order that d can be the smallest number to satisfy

$$F_n(x+d)-F_n(x-d)\geqslant k_n/n=n^{-(1+\delta)/(1+3\delta)}+O(n^{-1}),$$

d must satisfy

$$F(x+d)-F(x-d)=n^{-(1+\delta)/(1+3\delta)}+\theta_n n^{-(1+2\delta)/(1+3\delta)}c_n(\log n)^{1/2}\triangleq K_n^* \tag{9}$$

for n large, where $|\theta_n|\leqslant 1$.

Since $f\in\mathscr{F}_\delta$, there exists constant R such that

$$|f(x)-f(y)|\leqslant R|x-y|^\delta \text{ for any } x\in R^1,\ y\in R^1.$$

Hence

$$F(x+d)-F(x-d)=\int_{x-d}^{x+d}f(y)dy\begin{cases}\leqslant 2f(x)d+Gd^{1+\delta}\\ \geqslant 2f(x)d-Gd^{1+\delta}\end{cases} \tag{10}$$

with $G=2R(1+\delta)^{-1}$. Thus

$$d_1\leqslant a_n(x)\leqslant d_2, \tag{11}$$

where d_1. and d_2 are roots of Eqs. (12) and (13), respectively:

$$2f(x)d+Gd^{1+\delta}=K_n^*, \tag{12}$$

$$2f(x)d-Gd^{1+\delta}=K_n^*. \tag{13}$$

Define

$$S=\{x: x\in R^1;\ f(x)\leqslant Qn^{-\delta/(1+3\delta)}\}, \tag{14}$$

*) Here we tacitly make the assumption that

$$c_n=o((\log n)^{-1/2}n^{\delta/(1+3\delta)}).$$

Needless to say, this can be done without any loss of generality.

where Q is a constant independent of n to be chosen later. For $x\in S$, we have $f(x)>0$, and it follows from (12) that

$$d_1=\frac{K_n^*}{2f(x)}\left(1+\frac{G}{2f(x)}d_1^{\delta}\right)^{-1}. \tag{15}$$

Since $G>0$, it follows from (12) that $d_1\leqslant K_n^*/(2f(x))$. Hence from (15) we have

$$a_n(x)\geqslant d_1\geqslant\frac{K_n^*}{2f(x)}\left[1+\frac{G}{2f(x)}\left(\frac{K_n^*}{2f(x)}\right)^{\delta}\right]^{-1}\geqslant\frac{K_n^*}{2f(x)}\left[1-\frac{G}{2f(x)}\left(\frac{K_n^*}{2f(x)}\right)^{\delta}\right], \tag{16}$$

and we find, from (1) and (16), that

$$\hat{f}_n(x)\leqslant\frac{k_n}{nK_n^*}f(x)\left[1-\frac{G}{2f(x)}\left(\frac{K_n^*}{2f(x)}\right)^{\delta}\right]^{-1}. \tag{17}$$

Since

$$(2f(x))^{1+\delta}\geqslant(2Q)^{1+\delta}n^{-\delta(1+\delta)/(1+3\delta)},$$

$$GK_n^{*\delta}\leqslant 2^{\delta}Gn^{-\delta(1+\delta)/(1+3\delta)}.$$

Taking $Q=2G^{1/1+\delta)}$, we have

$$\frac{G}{2f(x)}\left(\frac{K_n^*}{2f(x)}\right)^{\delta}\leqslant\frac{1}{2}.$$

Also, $(1-x)^{-1}\leqslant 1+2x$ for $0\leqslant x\leqslant 1/2$. Hence by (17)

$$\hat{f}_n(x)\leqslant\frac{k_n}{nK_n^*}f(x)\left[1+\frac{G}{f(x)}\left(\frac{K_n^*}{2f(x)}\right)^{\delta}\right].$$

From this and the definition of k_n, K_n^*, it follows that

$$\hat{f}_n(x)-f(x)\leqslant 2c_n(\log n)^{1/2}n^{-\delta/(1+3\delta)}f(x)+\frac{k_n}{nK_n^*}G\left(\frac{K_n^*}{2f(x)}\right)^{\delta}. \tag{18}$$

From the fact that $f\in\mathscr{F}_\delta$ it follows that f is bounded on R^1. Also,

$$\lim_{n\to\infty}k_n/(nK_n^*)=1\quad\text{and}\ \frac{K_n^*}{f(x)}\leqslant\frac{2}{Q}n^{-1/(1+3\delta)}.$$

From (18) it follows that

$$\hat{f}_n(x)-f(x)=O(c_n(\log n)^{1/2}n^{-\delta/(1+3\delta)}) \tag{19}$$

uniformly for all x in S. On the other hand, writing

$$g(d)=2f(x)d-Gd^{1+\delta}-K_n^*,$$

we see that $g(K_n^*/2f(x))<0$, and

$$g(K_n^*/f(x))=K_n^*-G(K_n^*/f(x))^{1+\delta}=K_n^*[1-GK_n^{*\delta}/f^{1+\delta}(x)]. \tag{20}$$

Since

$$f^{1+\delta}(x)\geqslant Q^{1+\delta}n^{-\delta(1+\delta)/(1+3\delta)} \tag{21}$$

for $x\in S$,

$$GK_n^{*\delta}\leqslant G2^{\delta}n^{-\delta(1+\delta)/(1+3\delta)}\leqslant 2^{-1}Q^{1+\delta}n^{-\delta(1+\delta)/(1+3\delta)}. \tag{22}$$

From (20)—(22), we see that Eq. (13) has a root within the interval $[K_n^*/2f(x), K_n^*/f(x)]$. Hence $d_2 \leqslant K_n^*/f(x)$, and by (13) we get

$$d_2 = \frac{K_n^*}{2f(x)}\left[1 - \frac{G}{2f(x)} d_2^\delta\right]^{-1} \leqslant \frac{K_n^*}{2f(x)}\left[1 - \frac{G}{2f(x)}\left(\frac{K_n^*}{f(x)}\right)^\delta\right]^{-1}.$$

Using again $(1-x)^{-1} \leqslant 1+2x$ for $0 \leqslant x \leqslant 1/2$, it follows that

$$d_2 \leqslant \frac{K_n^*}{2f(x)}\left[1 + \frac{G}{f(x)}\left(\frac{K_n^*}{f(x)}\right)^\delta\right]$$

and from (1) we get

$$\hat{f}_n(x) \geqslant \frac{k_n}{nK_n^*} f(x)\left[1 + \frac{G}{f(x)}\left(\frac{K_n^*}{f(x)}\right)^\delta\right]^{-1} \geqslant \frac{k_n}{nK_n^*} f(x)\left[1 - \frac{G}{f(x)}\left(\frac{K_n^*}{f(x)}\right)^\delta\right].$$

Therefore

$$\hat{f}_n(x) - f(x) \geqslant -2c_n(\log n)^{1/2} n^{-\delta/(1+3\delta)} f(x) - \frac{k_n}{nK_n^*} G\left(\frac{K_n^*}{f(x)}\right)^\delta. \tag{23}$$

Similar to the deduction of (19) from (18), we get from (23)

$$\hat{f}_n(x) - f(x) \geqslant O(c_n(\log n)^{1/2} n^{-\delta/(1+3\delta)}) \tag{24}$$

uniformly for all x in S.

Now consider the case $x \bar{\in} S$. Choose $c > 0$ such that

$$2Qc + Gc^{1+\delta} = 1/2. \tag{25}$$

Note that the number c so determined is independent of both n and x (Q was defined previously as $2G^{1/(1+\delta)}$). As d_1 is the root of Eq. (12), for $x \bar{\in} S$ we have $d_1 \geqslant cn^{-1(1+3\delta)}$, since

$$\begin{aligned} &2f(x)cn^{-1/(1+3\delta)} + Gc^{1+\delta} n^{-(1+\delta)/(1+3\delta)} \\ &\quad \leqslant (2Qc + Gc^{1+\delta}) n^{-(1+\delta)/(1+3\delta)} = 2^{-1} n^{-(1+\delta)/(1+3\delta)} \end{aligned}$$

and by the definition of K_n^* it follows that $2^{-1} n^{-(1+\delta)/(1+3\delta)} < K_n^*$. This proves $d_1 \geqslant cn^{-1/(1+3\delta)}$, and $a_n(x) \geqslant d_1 \geqslant cn^{-1/(1+3\delta)}$. Hence by (1)

$$\hat{f}_n(x) \leqslant \frac{k_n}{2cn} n^{1/(1+3\delta)} \leqslant \frac{1}{2c} n^{-\delta/(1+3\delta)}.$$

Thus we have

$$|\hat{f}_n(x) - f(x)| \leqslant \hat{f}_n(x) + f(x) \leqslant \left(\frac{1}{2c} + Q\right) n^{-\delta/(1+3\delta)} \tag{26}$$

uniformly for all $x \bar{\in} S$. Finally, from (19), (24) and (26) we see that with probability one

$$\sup_x |\hat{f}_n(x) - f(x)| = O(c_n(\log n)^{1/2} n^{-\delta/(1+3\delta)}). \tag{27}$$

This proves the first part of the theorem.

For a proof of the second assertion of the theorem. we need the following result:

Lemma 2 Suppose that $f(x)\neq 0\neq f''(x)$ at some given point x, then for any $k_n\to\infty$ and $k_n=o(n^{4/5})$ we have

$$\sqrt{k_n}[\hat{f}_n(x)-f(x)]/f(x)\xrightarrow{L} N(0,1). \tag{28}$$

For a proof, see [3], Theorem 1.

Now take a density function $f\in\mathscr{F}_\delta$ satisfying the following conditions:

a. $f(x)=(1+\delta)|x|^\delta/2$, for $|x|$ sufficiently small,

b. There exists x_0 sonh that $f(x_0)\neq 0\neq f''(x_0)$,

c. There exists L such that $f(x)=0$ for $|x|\geqslant L$,

d. There exists u,v, $u<v$, such that $f(x)=1$ for $u<x<v$.

Let $\{k_n\}$ be any sequence of integers such that $1\leqslant k_n\leqslant n$ for $n=1,2,\cdots$. By choosing a subsequence if necessary, we can assume that $\{k_n\}$ satisfies one of the following conditions:

1. $k_n n^{-(1+2\delta)/(1+3\delta)}\to\infty$,
2. $k_n n^{-2\delta/(1+3\delta)}\to\infty$, but $k_n=O(n^{(1+2\delta)/(1+3\delta)})$,
3. $k_n\to\infty$, but $k_n=O(n^{2\delta/(1+3\delta)})$,
4. $k_n=k$ for n large, k is a positive integer.

Now we proceed to study these four cases separately:

Case 1. By virtue of condition c, we have $a_n(0)\leqslant L$, hence $\hat{f}_n(0)\geqslant k_n/(2Ln)$. Therefore

$$\begin{aligned} n^{\delta/(1+3\delta)}|\hat{f}_n(0)-f(0)| &\geqslant (2L)^{-1}n^{\delta/(1+3\delta)}n^{-1}k_n \\ &=(2L)^{-1}k_n n^{-(1+2\delta)/(1+3\delta)}\to\infty \end{aligned}$$

Case 2. Given $\varepsilon_1>0$. Take in lemma 1

$$B=2k_n/n, \qquad \varepsilon=\varepsilon_1\sqrt{k_n}\log n/n.$$

By the assumptions concerning k_n in this case, it is seen that $n\geqslant\max(B^{-1},8B\varepsilon^{-2})$, for n sufficiently large. Hence by lemma 1, for n sufficiently large,

$$P\left(\frac{n}{k_n^{1/2}\log n}\sup_x\{|\overline{F_n(b)-F_n(a)}-\overline{F(b)-F(a)}|:0\leqslant F(b)-F(a)\leqslant 2k_n/n\}\geqslant\varepsilon_1\right)$$

$$\leqslant 16n^2\exp\left(-\frac{\varepsilon_1^2 n^{-1}k_n(\log n)^2}{128k_n n^{-1}+4\varepsilon_1 k_n^{1/2}n^{-1}\log n}\right)+8n\exp\left(-\frac{k_n}{5}\right).$$

This quantity is of the order $o(n^{-2})$ by virtue of the assumptions of this case. Hence

$$\sup_x\{|\overline{F_n(b)-F_n(a)}-\overline{F(b)-F(a)}|:0\leqslant F(b)-F(a)\leqslant 2k_n/n\}=o(n^{-1}k_n^{1/2}\log n).\ a.s.$$

In the present case it is easily seen that $n^{-1}k_n^{1/2}\log n=o(k_n/n)$. Hence with probability one, in order that d can be the smallest number to satisfy

$$F_n(d)-F_n(-d)\geqslant k_n/n$$

for n large, d must satisfy

$$F(d)-F(-d)=k_n/n+\theta_n n^{-1}k_n^{1/2}\log n, \qquad |\theta_n|\leqslant 1. \tag{29}$$

From this and the assumption on k_n in the present case, and note the condition a, one sees that d is small when n is large. Hence for these n

$$F(d)-F(-d)=\int_{-d}^{d}f(y)dy=d^{1+\delta}.$$

Therefore d is the solution of the equation

$$d^{1+\delta}=k_n/n+\theta_n n^{-1}k_n^{1/2}\log n$$

and for n sufficiently large,

$$a_n(0)=d\leqslant(2n_n/n)^{1/(1+\delta)}.$$

Hence it follows from (1) that

$$\hat{f}_n(0)=\frac{k_n}{2a_n(0)n}\geqslant\frac{1}{4}\left(\frac{k_n}{n}\right)^{\delta/(1+\delta)}.$$

Since $k_n n^{-2\delta/(1+3\delta)}\to\infty$, we have

$$n^{\delta/(1+3\delta)}|\hat{f}_n(0)-f(0)|=n^{\delta/1+3\delta)}\hat{f}_n(0)\to\infty.$$

Case 3. Since $k_n\to\infty, k_n=O(n^{2\delta/(1+3\delta)})$ and $2\delta/(1+3\delta)<4/5$ for $0<\delta\leqslant 1$, lemma 2 can be employed at point x_0, and $k_n^{1/2}(\hat{f}_n(x_0)-f(x_0))/f(x_0)\xrightarrow{L}N(0,1)$. Since the support set of $N(0,1)$ is R^1, $\hat{f}_n(x_0)-f(x_0)$ cannot, with probability one, have an order $O(k_n^{-1/2})=O(n^{-\delta/(1+3\delta)})$.

Case 4. Choose arbitrarily $y\in(u,v)$. Define

$$Y_n=\#(\{i:1\leqslant i\leqslant n;\ |X_i-y|\leqslant k/(4n)\}).$$

Then by condition d and the well-known fact of approximating binomial distribution by Poisson distribution, one sees easily that

$$\lim_{n\to\infty}P(Y_n=k)=e^{-k/2}(k/2)^k/k!\triangleq A>0. \tag{30}$$

Since $P(a_n(y)\leqslant k/(4n))\geqslant P(Y_n=k)$, one sees from (1) that, with probability not less than $P(Y_n=k)$, $\hat{f}_n(y)\geqslant k/\{2n(k/4n)\}=2$. Hence it follows from (30) that for n sufficiently large we have

$$|\hat{f}_n(y)-f(y)|\geqslant 2-1=1$$

with probability not less than $A/2>0$.

Summing up the above discussion. We see that in no case is the assertion

$$\sup_x|\hat{f}_n(x)-f(x)|=O(n^{-\delta/(1+3\delta)}), \qquad a.s.$$

true. This concludes the proof of the theorem.

Remark. The method of this paper can equally be employed in the case of high-dimensional densities.

References

[1] Loftsgarden, D. O. and Quesenberry, C. F., A non-parametric estimate of a multivariate density function, Ann. Math. statist. 36 (1965), 1049—1051.

[2] Devroye, L. P. and Wagner. T. J., The strong uniform consistency of nearest neighbor density estimates, Ann. Statist. 5 (1977), 536—540.

[3] Chen Xiru (陈希孺), Convergence Rates of Nearest Neighbor Density Estimates, Scientia Sinica, 1982.

[4] Devroye, L. P. and Wagner, T. J., The strong uniform consistency of kernel density estimates, Proc. of the fifth inter. symp. multi. analysis (1980), 59—77.

最近邻密度估计的一致收敛速度

陈希孺

中 文 摘 要

设 $X_1, \cdots X_n$ 是从具密度函数 f 的一维总体中抽出的 iid. 样本。1965 年，Loftsgarden 等在 [1] 中提出了如下的估计 $f(x)$ 的方法：选择最小的 $a_n(x)=a_n(x; X_1, \cdots, X_n)$，使区间 $[x-a_n(x),\ x+a_n(x)]$ 中至少包含 $X_1, \cdots, X_n$ 中的 k_n 个样本。此处 k_n 为一适当选择的整数，$1 \leqslant k_n \leqslant n$。然后以 $\hat{f}_n(x)=k_n/\{2na_n(x)\}$ 作为 $f(x)$ 的估计。这种估计通常称为“最近邻估计”。有一些作者研究了这种估计的相合性。本文作者在 [3] 中研究了这种估计的一致强收敛速度，得出了初步结果，在本文中，我们显著地改进了上述结果：

定理 设 $0<\delta\leqslant 1$。以 $\mathscr{F}_\delta$ 记所有满足 δ 阶 Lipshitz 条件的一维概率密度函数的族。若取

$$k_n=[n^{2\delta/(1+3\delta)}]$$

则对任何常数 $c_n \to \infty$ 有

$$\sup_x |\hat{f}_n(x)-f(x)| = O(n^{-\delta/(1+3\delta)}(\log n)^{1/2} c_n),\ a.s.$$

对任何 $f\in\mathscr{F}_\delta$。另一方面，对任何 δ，$0<\delta\leqslant 1$，可找到 $f\in\mathscr{F}_\delta$，使不论怎样选择 k_n，下述断言

$$\sup_x |\hat{f}_n(x)-f(x)| = O(n^{-\delta/(1+3\delta)}),\ a.s.$$

都不可能成立。

本定理说明：对满足 δ 阶 Lipshitz 条件的概率密度函数族的全体而言，收敛速度的主要部分即 $n^{-\delta/(1+3\delta)}$ 中的指数 $\delta/(1+3\delta)$ 已无可改进。当 $\delta=1$ 时，这个值是 1/4。这个结论曾由作者在 [3] 中作为一个猜测提出来过。

又，本文方法对处理多维密度的最近邻估计也适用。