第19卷 第2期　　　　郑州工业大学学报　　　　Vol.19 No.2
1998年 6 月　　Journal of Zhengzhou University of Technology　　Jun. 1998

# 不连续密度函数NN—估计的收敛速度

刘妍岩　　　　　　赵玲玲
(武汉大学数学系,武汉,430072)　(河南周口师专数学系,466000)
王霞
(郑州工业大学数力系)

**摘　要**　讨论不连续密度函数最近邻估计(NN—估计)的收敛速度。由于密度函数不连续时的NN—估计不是相合的,我们提出一个改进的NN—估计,使之有相合性,且收敛速度不变。

**关键词**　NN—估计;α阶拟光滑函数;核函数
**中图分类号**　O174

## 1　定义及引理

设 $x_1,\cdots,x_n$ 为从某个具有分布函数 $F$ 和密度函数 $f$ 的一维总体中抽出的独立同分布 $(i,i,d)$ 样本,选定一个与 $n$ 有关的自然数 $k_n<n$,找出最小的 $a_n(x)$,使区间 $[x-a_n(x), x+a_n(x)]$ 至少含有 $k_n$ 个样本点,定义 $f(x)$ 的最近邻 估计($NN$—估计)如下:

定义1.1:称 $f_n(x)=(n a_n(x))^{-1}\sum_{i=1}^{n}k\left(\dfrac{x_i-x}{a_n(x)}\right)$ 为密度函数 $f(x)$ 在 $x$ 点的最近邻估计,其中 $k(x)$ 为核函数。

密度函数 $f_n(x)$ 的收敛速度结果参看文献[1],[2],[3],[4],[5]。

定义1.2:函数 $f(x)$ 在 $x$ 点处称为α阶拟光滑的,若:

(1)存在开区间 $J$ 且 $x\in J$.使 $f^{(m)}(t)$ 在 $J\setminus\{x\}$ 上存在,连续且有界($m$ 表示小于α的最大整数),进而,$f^{(m)}(x+0)$ 和 $f^{(m)}(x-0)$ 存在;

(2) $f^{(i)}(x+0)=f^{(i)}(x-0)$, $i=1,\cdots,m$ 若 $\alpha>1$;

(3)极限 $\lim\limits_{z\to 0^+}|z|^{-\alpha}[f(x+z)-\sum_{i=0}^{m}f^{(i)}(x+0)\dfrac{z^i}{i!}]=f_\alpha^+(x)$,

$\lim\limits_{z\to 0^+}|z|^{-\alpha}[f(x+z)-\sum_{i=0}^{m}f^{(i)}(x-0)\dfrac{z^i}{i!}]=f_\alpha^-(x)$ 存在.

以下证 $f_a^*(x)=f(x+0)+(1-a)f(x-0)$, $0<a<1$,

特别,证 $f^*(x)\triangleq f_{\frac{1}{2}}^*(x)=\dfrac{1}{2}[f(x+0)+f(x-0)]$;

证 $f_\alpha(x)=f_\alpha^+(x)+f_\alpha^-(x)$,文中所出现 $a_n(x)$,$k_n$ 等定义同定义1.1,不再重述。

首先,我们需要下述引理:

收稿日期:1997—11—24
第一作者　女　1967年5月生　硕士学位　讲师

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://w

引理 1.1[6]：设随机变量 $Y$ 服从二项分布 $B(n, p)$，则对任意 $\varepsilon>0$，

有
$$P\left\{\left|\frac{Y}{n}-p\right| \geqslant \varepsilon\right\} \leqslant \alpha \cdot \exp\left(-\frac{n\varepsilon^{2}}{2\bar{p}+\varepsilon}\right)$$

其中
$$\bar{p}=\min(p, 1-p)$$

引理 1.2：若密度函数 $f(x)$ 在 $x_0$ 处满足：$f_0^{*}(x)>0$ 且存在 $\delta>0$，使得：
$$|f(x)-f(x_0+0)| \leqslant R|x-x_0| \qquad \forall\ x \in (x_0, x_0+\delta);$$
$$|f(x)-f(x_0-0)| \leqslant R|x-x_0| \qquad \forall\ x \in (x_0-\delta, x_0)$$

则当 $\frac{k_n}{n} \to 0, \frac{k_n}{2\log n} \to 0\ (n \to \infty)$ 时，对任一数串 $C_n \to \infty$（不失一般性可设 $C_n \frac{k_n}{n} \to 0$）

有
$$\lim_{n \to \infty} \sup\ a_n(x_0) \frac{n}{k_n} C_n^{-1} \to 0 \quad a \cdot s \cdot \tag{1}$$
$$\lim_{n \to \infty} a_n(x_0) \frac{n}{k_n} C_n^{-1} \to \infty \quad a \cdot s \cdot \tag{2}$$

证明：我们只证明(1)，(2)的证明类似：

证：$N_n(x, a)=\#\{i: 1 \leqslant i \leqslant n, x-a \leqslant x_i \leqslant x+a\}$ $\#\{\cdots\}$表示集合元素的个数。

对任意 $\varepsilon>0$，
$$P\left(\frac{n}{k_n} C_n^{-1} a_n(x_0) \geqslant \varepsilon\right)$$
$$=P\left(N_n\left(x_0, \varepsilon \frac{c_n k_n}{n}\right) \leqslant k_n\right)$$
$$=P\left(\frac{N_n\left(x_0, \varepsilon \frac{c_n k_n}{n}\right)}{n}-P_n \leqslant \frac{k_n}{n}-P_n\right)$$

其中
$$p_n=\int_{x_0-\varepsilon \frac{c_n k_n}{n}}^{x_0+\varepsilon \frac{c_n k_n}{n}} f(t)\, dt$$

易知，存在常数
$$c>0, \text{使}\ p_n \geqslant \varepsilon \cdot c \cdot c_n \frac{k_n}{n} \tag{3}$$

由引理 1.1 及(3)知：
$$P\left(\frac{n}{k_n} C_n^{-1} a_n(x_0) \geqslant \varepsilon\right)$$
$$\leqslant P\left\{\left|\frac{N_n\left(x_0, \varepsilon \frac{c_n k_n}{n}\right)}{n}-p_n\right| \geqslant p_n-\frac{k_n}{n}\right\} \quad .1$$
$$\leqslant \alpha \cdot \exp\left\{-\frac{c}{32} \varepsilon \cdot c_n \cdot k_n\right\}$$
$$\leqslant \alpha \cdot \exp\left\{-\frac{c}{32} \varepsilon k_n\right\}$$

由 Borel—Canteli 引理知(1)式成立。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://w

证：$\mu$ 及 $\mu_n$ 分别为分布函数 $F$ 及 $F$ 的经验分布函数 $F_n$ 诱导的测度；[a] 表示不超过 a 的最大整数。

引理 1.3：设核函数 $k(x)$ 满足下述条件：

$(a)$：$k(x)$ 为 $R'$ 上有界变差函数；

$(b)$：$|x|>\rho(\rho>0)$ 时，$k(x)\equiv 0$，且 $k(x)$ 在 $[-\rho,\rho]$ 上连续；

$(c)$：$k(\rho)$，$k(0)$，$k(-\rho)$ 都是大于 0 且它们相等；若密度函数 $f(x)$ 在 $x$ 处为 $r$ 阶 $(r>1)$ 拟光滑的，则当 $k_n=[n^{\frac{2r}{2r+1}}(\log n)^{\frac{1}{2r+1}}]$ 时，对任一数串 $c_n\to\infty^{+}$

有
$$\lim_{n\to\infty}\sup\left(\frac{n}{\log n}\right)^{\frac{1}{2r+1}}C_n^{-1}|f_n(x)-\tilde f_n(x)|\to 0\quad a.s.$$

其中
$$\tilde f_n(x)=a_n^{-1}(x)\int_{0<\left|\frac{y-x}{a_n(x)}\right|<\rho}k\left(\frac{y-x}{a_n(x)}\right)u(dy)$$

证明：不失一般性，设 $0<\rho<1$，由引理 1.2 知，当 $n$ 充分大时，以概率 1 有

$$\begin{aligned}
&|f_n(x)-\tilde f_n(x)|\\
&=a_n^{-1}(x)\left|\int_{0<z\leq\rho}k(z)\mu_n(nd(x+a_n(x)z)-\int_{0<z\leq\rho}k(z)\mu_n(d(x+a_n(x)z)\right.\\
&\left.+\int_{-\rho\leq z<0}k(z)\mu_n(d(x+a_n(x)z)-\int_{-\rho\leq z<0}k(z)\mu_n(d(x+a_n(x)z)\right|\\
&\leq a_n^{-1}(x)k(\rho)\{|\mu_n[(x,x+a_n(x)\rho)]-\mu[(x,x+a_n(x)\rho)]\\
&+\mu_n[(x-a_n(x)\rho,x)]-\mu[(x-a_n(x)\rho,x)]|\\
&+\left|\int_{0<z\leq\rho}[\mu_n[(x,x+a_n(x)z)]-\mu[(x,x+a_n(x)z)]\;k(dz)\right.\\
&\left.+\int_{-\rho\leq z<0}[\mu_n[(x+a_n(x)z,x)]-\mu[(x+a_n(x)z,x)]]k(dz)\right|\}\\
&\leq S\cdot\sup_{A_n}|\mu_n(A)-\mu(A)|\cdot\frac{1}{a_n(x)}\\
&\leq S\cdot\frac{n}{k_n}C_n^{\frac{1}{2}}\sup_{A_n}|\mu_n(A)-\mu(A)|
\end{aligned}\tag{4}$$

其中 $A_n$ 为 $R'$ 中所有满足 $\mu(A)\leq C_n^{\frac{1}{3}}\left(\frac{\log n}{n}\right)^{\frac{1}{2r+1}}$ 的区间集，$S$ 代表常数。

类似[7]的证明可得：

$$\lim_{n\to\infty}\sup\left(\frac{n}{\log n}\right)^{\frac{r+1}{2r+1}}C_n^{-\frac{2}{3}}\sup_{A_n}|\mu_n(A)-\mu(A)|=0\quad a.s.$$

结合(4)可知引理的结论成立。

## 2　$f_n(x)$的收敛速度

定理 2.1：设密度函数 $f(x)$ 在 $x$ 处为 $r$ 阶拟光滑的$(r>1)$，且 $f^*(x)>0$，设核函数 $k$ 满足引理 1.3 的条件，且：

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://w

$$(1)\int_{-\rho}^{\rho}k(x)dx=1,\int_{0}^{\rho}k(x)dx=a,0<a<1$$

$$(2)\int_{-\rho}^{\rho}x^{i}k(x)dx=0,i=1,\cdots,m$$

$m$ 为小于 $r$ 的最大整数，则当 $k_n=[n^{\frac{2r}{2r+1}}(\log n)^{\frac{1}{2r+1}}]$ 时，对任意一数串 $C_n\to\infty(n\to\infty)$，有：

$$\lim_{n\to\infty}\sup\left(\frac{n}{\log n}\right)^{\frac{1}{2r+1}}C_n^{-1}|f_n(x)-f_a^*(x)|=0\quad a.s.$$

证明：由引理 1.3 知，为证本定理，只须证明：

$$\lim_{n\to\infty}\sup\left(\frac{n}{\log n}\right)^{\frac{1}{2r+1}}C_n^{-1}|\tilde{f}_n(x)-f_a^*(x)|=0\quad a.s.\tag{5}$$

$\tilde{f}_n(x)$定义同引理 1.3。

由引理 1.2 知，$n$ 充分大时，以概率 1 有：

$$|\tilde{f}_n(x)-f_a^*(x)|$$

$$=a_n^{-1}\int_{\left|\frac{y-x}{a_n(x)}\right|\leq\rho}k\left(\frac{y-x}{a_n(x)}\right)f(y)dy-f_a^*(x)$$

$$=\left|\int_0^{\rho}k(z)\left\{\sum_{i=1}^{m}f^{(i)}(x+0)\frac{(a_n(x)z)^i}{i!}+[0(1)+f_r^+(x)]\right\}(a_n(x)z)^r dz\right.$$

$$=\int_{-\rho}^{0}k(z)\left\{\sum_{i=1}^{m}f^{(i)}(x-0)\frac{(a_n(x)z)^i}{i!}+[0(1)+f_r^-(x)]\right\}(a_n(x)z)^r dz$$

$$\leq 3a_n^r(x)\max\{|f_r^+(x)|,|f_r^-(x)|\}\cdot\int_0^{\rho}|k(z)|z^r dz$$

$$\leq S\cdot\left(\frac{\log n}{n}\right)^{\frac{r}{2r+1}}C_n^{\frac{r}{2r+1}}\quad a.s.\quad(\ )$$

由此可知(5)式成立，因此定理得证。

注：(1)当 $k(x)$为关于原点对称的密度函数时，定理 2.1 的条件(2)可改为：$f^{(2i)}(x+0)=f^{(2i)}(x-0)=0,i=1,\cdots,\left[\frac{m}{2}\right]$，($m$ 为小于 $r$ 的最大整数)，在实际应用中，常遇到这种情况。

(2)由定理 2.1 的结论可知，一般情况下 $f_n(x)$不是 $f(x)$的相合估计，在第 3 节中我们将对此加以改进。

下面给出比定理 2.1 更一般的结论：

以下设 $f(x)$为 $r(r>1)$阶拟光滑的，记：

$$f_{(j)}^*(x)\triangleq\frac{1}{2}(f^{(i)}(x+0)+f^{(i)}(x-0))\quad i=1,\cdots,m$$

($m$ 为小于 $r$ 的最大整数)，设 $f^*(x)>0,k_n=O(n)$，构造函数 $g_n(x)$如下：

$$g_n(x)=f^*(x)+\sum_{j=1}^{l}h_j(x)\left(\frac{k_n}{2n}\right)^{2j}+O\left(\left(\frac{k_n}{n}\right)^{2l+2}\right)$$

其中 $h_j$ 仅依赖于$f^*(x),f_{(2)}^*(x),\cdots,f_{(2j)}^*(x)$，且当 $f_{(2)}^*(x)=\cdots=f_{(2j-2)}^*(x)=0$ 时，

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://w

$$h_j(x)=\frac{f^*_{(2j)}(x)}{(2j+1)!\,(f(x))^{2j}}$$

而 $f^*_{(2j)}(x)=0$ 时，必有 $h_j(x)=0$。

定理 2.2：设密度函数 $f(x)$ 在 $x$ 处为 $r(r\geqslant 2)$ 阶拟光滑的，且 $f^*(x)>0$，设核函数 $k(x)$ 满足：

(1) $k(x)$ 为 $R'$ 上具有有界变差的概率密度函数，且关于原点对称；

(2) $|x|>\rho(\rho>0)$ 时，$k(x)\equiv 0$；且 $k(x)$ 在 $[-\rho,\rho]$ 上连续；

(3) $\int_{-\rho}^{\rho}k(x)x^{2i}dx=\frac{1}{2i+1}$，$i=1\cdots,l$，$l=\left[\frac{r}{2}\right]$，则当 $k_n=[n^{\frac{2r}{2r+1}}(\log n)^{\frac{1}{2r+1}}$ 时，对任一数串 $c_n\to\infty$，$(n\to\infty)$ 有：

$$\lim_{n\to\infty}\sup\left(\frac{n}{\log n}\right)^{\frac{r}{2r+1}}C_n^{-1}|f_n(x)-g_n(x)|=0\quad a.s.$$

证明：类似[4]的证明，易知：

$$\lim_{n\to\infty}\sup\left(\frac{n}{\log n}\right)^{\frac{r}{2r+1}}e_n^{-1}|f_n(x)-f_n(x)|=0\quad a.s.$$

下面只须证明：

$$\lim_{n\to\infty}\sup\left(\frac{n}{\log n}\right)^{\frac{r}{2r+1}}C_n^{-1}|f_n(x)-g_n(x)|=0\quad a.s.$$

($f_n(x)$ 见引理 1.3)。首先，证明

$$\lim_{n\to\infty}\sup\left(\frac{n}{\log n}\right)^{\frac{r}{2r+1}}C_n^{-1}\left|a_n(x)-\frac{k_n}{2n\cdot g_n(x)}\right|=0\quad a.s.\tag{6}$$

$\forall\ \varepsilon>0$，

$$\begin{aligned}&P\left\{\left(\frac{n}{\log n}\right)^{\frac{r}{2r+1}}C_n^{-1}\left(a_n(x)-\frac{k_n}{2n\cdot g_n(x)}\right)\geqslant\varepsilon\right\}\\&=P\left\{a_n(x)\geqslant d_n\right\}\\&=P\left\{\frac{N_n(x,d_n)}{n}-p_n\leq\frac{k_n}{n}-p_n\right.\end{aligned}\tag{7}$$

其中 $d_n\triangleq\frac{k_n}{2n\cdot g_n(x)}+C_n\varepsilon\left(\frac{\log n}{n}\right)^{\frac{r}{2r+1}}$，$p_n=\int_{x-d_n}^{x+d_n}f(t)dt$

不妨设 $C_n\left(\frac{\log n}{n}\right)^{\frac{r}{2r+1}}\to 0(n\to\infty)$，则

$$\begin{aligned}\rho_n&=\int_{x-d_n}^{x+d_n}f(t)dt\\&=2f^*(x)d_n+\sum_{i=1}^{l}\frac{f^*_{(2i)}(x)}{(2i+1)!}d_n^{2i+1}+\frac{f_r(x)}{r+1}d_n^{r+1}+O(d_n^{r+1})\\&=2d_n\left\{f^*(x)+\sum_{i=1}^{l}\frac{f^*_{(2i)}(x)}{(2i+1)!}d_n^{2i}+\frac{f_r(x)}{r+1}d_n^{r}\right\}+O(d_n^{r+1})\\&=\frac{k_n}{n\cdot g_n(x)}\left\{f^*(x)+\sum_{i=1}^{l}\left[\frac{f^*_{(2i)}(x)}{(2i+1)!}\cdot\left(\frac{k_n}{2n\cdot g_n(x)}\right)^{2i}\right.\right.\end{aligned}$$

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://w

$$+2\varepsilon f^{*}(x)C_n\left(\frac{\log n}{n}\right)^{\frac{r}{2r+1}}+O\left(\left(\frac{\log n}{n}\right)^{\frac{r}{2r+1}}\right) \tag{8}$$

由引理 1.1 知：

$$(7)\text{式}\le P_r\left\{\left|\frac{N_n(x,d_n)}{n}-p_n\right|\geqslant \varepsilon f^{*}(x)C_n\left(\frac{\log n}{n}\right)^{\frac{r}{2r+1}}\right\}$$

$$\le 2\exp\left\{\frac{-n(\varepsilon f^{*}(x)C_n)^2\left(\frac{\log n}{n}\right)^{\frac{r}{2r+1}}}{6\left(\frac{\log n}{n}\right)^{\frac{r}{2r+1}}\varepsilon f^{*}(x)C_n}\right\}$$

$$\le 2\exp\left\{-n^{\frac{1}{2r+1}}\right\} \tag{9}$$

(9)式右端为收敛级数的通项，由 Borel—Canteli 引理知：

$$\limsup_{n\to\infty}\left(\frac{n}{\log n}\right)^{\frac{r}{2r+1}}C_n^{-1}\left(a_n(x)-\frac{k_n}{2n\cdot g_n(x)}\right)\le \varepsilon \quad a\cdot s\cdot$$

类似可证：

$$\liminf_{n\to\infty}\left(\frac{n}{\log n}\right)^{\frac{r}{2r+1}}C_n^{-1}\left(a_n(x)-\frac{k_n}{2n\cdot g_n(x)}\right)\geq -\varepsilon \quad a\cdot s\cdot$$

由 ε 的任意性，知(6)式成立。

其次，考虑：

$$\left|\hat{f}_n(x)-g_n(x)\right|$$

$$=\left|\int_{|z|\le\rho}k(z)f(x+a_n(x)z)dz-g_n(x)\right|$$

$$=\left|\int_0^{\rho}f(x+0)k(z)dz+\sum_{i=1}^{m}\int_0^{\rho}k(z)\frac{z^i}{i!}f^{(i)}(x+)a_n^i(x)dz\right.$$

$$+\sum_{i=0}^{m}\int_{-\rho}^{0}k(z)\frac{f_i(x-)}{i!}(a_n(x)z)^i dz+\int_0^{\rho}k(z)f_r^{+}(x)(a_n(x)z)^r dz$$

$$\left.+\int_{-\rho}^{0}k(z)f_r^{-}(x)(-a_n(x)z)^r dz+O(1)\cdot\int_{-\rho}^{\rho}(a_n(x)z)^r k(z)dz-g_n(x)\right|$$

$$=\sum_{i=1}^{l}\frac{f_{(2i)}^{*}(x)}{(2i+1)!}\left\{(a_n(x))^{2i}-\left(\frac{k_n}{2n\cdot g_n(x)}\right)^{2i}\right\}+O(1)\cdot a_n^r(x)+O\left(\left(\frac{k_n}{n}\right)^{2l+2}\right)$$

$$\le\sum_{i=1}^{l}\left\{\frac{f_{(2i)}^{*}(x)}{(2i+1)!}\left|a_n(x)-\frac{k_n}{2n\cdot g_n(x)}\right|\cdot\left|\sum_{j=0}^{2i}a_n^j(x)\left(\frac{k_n}{2n\cdot g_n(x)}\right)^{2i-1-j}\right|\right\}+O(a_n^r(x))$$

由(6)知：

$$\limsup_{n\to\infty}\left(\frac{n}{\log n}\right)^{\frac{r}{2r+1}}C_n^{-1}\left|\hat{f}_n(x)-g_n(x)\right|=0 \quad a\cdot s\cdot$$

因此，定理结论成立。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://w

## 3　改进的 $NN$—估计

现在我们对密度函数 $r$ 阶拟光滑点处的 $NN$ —估计加以改进，使之成为相合估计。

首先，给出下述定义：

定义 3.1：令 $k_n$ 为一与 $n$ 有关的自然数，$k_n \leqslant \frac{n}{2}$，定义 $b_n(x)$为最小的 $b_n$，使区间$(x, x+b_n)$中至少包含样本总体 $x_1, \cdots, x_n$ 的$k_n$ 个样本点，$C_n(x)$ 为最小的 $C_n$，使区间$(x-C_n, x$ 中至少包含样本总体的$k_n$ 个样本点，称：

$$f_{n1}(x) = (nb_n(x))^{-1}\sum_{x_i > x} k\left(\frac{x_i - x}{b_n(x)}\right) \cdot \frac{1}{a} \qquad 及$$

$$f_{n2}(x) = (nC_n(x))^{-1}\sum_{x_i < x} k\left(\frac{x_i - x}{C_n(x)}\right) \cdot \frac{1}{1-a}$$

为 $f(x)$在 $x$ 处的右最邻估计及左最近邻估计。

其中 $a = \int_0^{+\infty} k(x)\,dx \in (0,1)$.

定义 3.2：称函数 $f(x)$在 $x$ 点处为 $\alpha$ 阶右拟光滑的($\alpha > 0$)，若：

(1)存在 $b > 0$，使 $f^{(i)}(t)$在区间$(x, x+b)$存在，连续有界，$i=1, \cdots, m$($m$ 为小于 $\alpha$ 的最大整数)；

(2)极限 $\lim\limits_{z \to 0^+} |z|^{-\alpha}\left(f(x+z) - \sum\limits_{i=0}^{m} f^{(i)}(x+0)\frac{z^i}{i!}\right) = f_\alpha^{+}(x)$ 存在。类似可定义 $\alpha$ 阶左拟光滑。

类似引理 1.2 有：

引理 3.1：若 $f(x_0+0) > 0$，且存在 $z > 0$，使：

$$|f(x) - f(x_0+0)| \leqslant R|x - x_0| \qquad x \in (x_0, x_0+\delta),$$

则$\frac{k_n}{n} \to 0$, $\frac{k_n}{2\log n} \to \infty$ ($n \to \infty$)时，对任一数串 $C_n \to \infty$，(不妨设 $C_n \frac{k_n}{n} \to 0$ $(n \to \infty)$有：

$$\limsup_{n \to \infty} b_n(x) \frac{n}{k_n} C_n^{-1} = 0 \qquad a \cdot s \cdot$$

$$\liminf_{n \to \infty} b_n(x) \frac{k_n}{n} C_n = 0 \qquad a \cdot s \cdot$$

对 $C_n(x)$有类似的结论。

类似 $Th2.2$，有下面结论：

定理 3.1：设密度函数 $f(x)$在 $x = x_0$ 处在 $r$ 降($r > 1$)右拟光滑的，核函数 $k(x)$满足：

$(a)$：$k(x)$为 $R'$上具有有界变差的概率密度函数；

$(b)$：当$|x| > \rho$($\rho > 0$) 时，$k(x) \equiv 0$；且 $k(x)$在$(-\rho, \rho)$上连续：

$(c)$：$f^i(x_0+) = 0$，$i = 1, \cdots, m$，($m$ 为小于 $r$ 的最大整数)，则当 $k_n = [n^{\frac{2r}{2r+1}}(\log n)^{\frac{1}{2r+1}}$ 时，对任一数串 $C_n \to \infty$，有：

$$\limsup_{n \to \infty}\left(\frac{n}{\log n}\right)^{\frac{r}{2r+1}} C_n^{-1} |f_{n1}(x_0) - f(x_0+)| = 0 \qquad a \cdot s \cdot$$

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://w

注：(1)对 $f_{n2}(x)$有类似的结论；

(2)特别，取 $C_n=\left(\frac{n}{\log n}\right)^{\frac{r}{2r+1}}$时，可知：

$$\lim_{n\to\infty}\left|f_{n1}(x_0)-f(x_0+)\right|=0$$

则 $f_{n1}(x_0)$为 $f(x_0+)$的相合估计，与 $Th2.2$ 比较，易知 $f_{n1}$的收敛速度与 $f_n(x)$相同。

## 参考文献


1 陈希孺.最近邻密度估计的收敛速度.中国科学，12(1981)，1419～1428

2 陈希孺.最近邻密度估计的渐近均方误差.数学年刊，12(1981)，425～429

3 白志东.中国科技大学学报，5(1983)No.1，114～118

4 E.l：ebcher：kenel Estimators for probability Densities with Discontinuities statistics 21(1990)，No.2

5 Daren.B.Chine & Jefferey.Hart：kenel Estimators of Densities with Ciscontinuities or Discontinuous Derivatives，statistics 22(1991)，69～84

6 Hoeffding.W：J.Amer.Statist.Assoo.58(1963)，13～30

7 卢江.最近邻密度估计的逐点强收敛速度.应用概率统计，2(1986)No.1.21～27


# The Corvergence rate of NN—estimation for Discontinuous Density Function


Liu Yanyan Zhao Lingling

(*Wuhan University*)(*Henan Zhoukou Teachers' College*)

Wang Xia

(*Zhengzhou University of Technology*)



**Abstract** This paper discusses the convergence rate of kernel estimators (NN—estimation) for discontinuous function. Because NN—estimation is not congruent when density function is discontinuous, we give an improved NN—estimation so that it has congruence and the same convergence rate.



**Keywords** NN—estimation；α—order quasi—function；kernel function


(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://w