1982年 第2卷 第4期
Vol.2, No.4, 1982

数学研究与评论
JOURNAL OF MATHEMATICAL RESEARCH & EXPOSITION

Rosenblatt 估计一致收敛的必要条件*

陈希孺

（中国科技大学）

§1 引言和结果

设 $X_1, \cdots, X_n$ 为随机变量 X 的独立同分布观察值．在以下，F 专用于记 X 的分布函数．若 X 有概率密度函数，则总用 f 来记．通过样本 $X_1, \cdots, X_n$ 去估计 f，是一个重要的统计问题．到如今积累了大量的文献．其中最重要的一类估计是所谓"核估计"，其发端是 M. Rosenblatt 的工作[1]：取一串数 $\{h_n\}$，$0<h_n \to 0$．以 $N_n(a,b)$ 记样本 $X_1, \cdots, X_n$ 中落在 $[a,b)$ 内的个数，而令

$$f_n(x) = N_n(x-h_n,\ x+h_n)/(2nh_n) \tag{1}$$

Rosenblatt 取 f_n 为 f 的估计．引进函数

$$K(x) = \begin{cases} 1/2, & \text{当} -1<x\leqslant 1, \\ 0, & \text{其他 } x \end{cases} \tag{2}$$

则可将估计(1)改写为：

$$f_n(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right) \tag{3}$$

E. Parzen 在[2]中，将上述估计推广为广泛的一类估计，其中 K 不必有（2）的形式，而有相当的选择自由，这类估计叫核估计，而 K 称为（该估计的）核函数．

核估计的大样本性质是文献中的一个热门题目．1962 年，Parzen 在[2]中证明了：若 $h_n \to 0$，$nh_n^2 \to \infty$，f 在 R^1 上一致连续，而核函数 K 满足一些条件，则当 $n\to\infty$ 时 $\sup_x |f_n(x)-f(x)| \xrightarrow{P} 0$（依概率收敛）．Nadaraya 1965 年在[3]中证明了：当 f 在 R^1 一致连续而 K 满足一定条件时，有 $\sup_x |f_n(x)-f(x)| \to 0$，$a, e$．以后陆续发表了一些结果．1980 年 Devroye 等在[4]中对以前的结果作了较大的改进．

1969 年，Schuster 在[5]中讨论了反面的问题：简言之，即在上引结果中，"f 在 R^1

* 1981 年 9 月 7 日收到．

上一致连续”的条件是否必要．Schuster 在一定程度上回答了这个问题，他证明了：若 h_n 和核函数 K 满足一定的条件，则如存在函数 g，使

$$\sup_x |f_n(x)-g(x)|\to 0 \quad a.e.,\ 当\ n\to\infty \tag{4}$$

则 g 必在 R^1 上一致连续，且就是 X 的密度函数．

Schuster 加在 h_n 和 K 上的条件较强．例如，要求函数 K 在 R^1 上连续有界变差．这就把重要的 Rosenblatt 估计排除在外．我们对这个特例作了研究，得到了以下比较彻底的结果．

定理 沿用前面的记号，设 $\{h_n\}$ 为一串数，$0<h_n\to 0$，f_n 由(1)式定义．则如存在定义于 R^1 上的函数 g，使(4)式成立，则 g 必在 R^1 上一致连续，且 X 有密度 $f=g$．

值得注意的是，上述定理对 $\{h_n\}$ 没有作什么特殊的要求，我们甚至可证明，$h_n\to 0$ 这要求也是必要的．当然，我们可以进一步提出一个一般性的问题：为了 Schuster 的结论成立，加在 h_n 和 K 上的条件可以减轻到何种程度．上述定理只是对这个问题迈出的很小的一步．但这个结果有一定的兴趣．一则由于 Rosenblatt 估计在核估计类中是最重要的．一则由于核函数(2)在非连续核函数中有一定的代表性，例如，上述定理使我们有根据猜想：Schuster 的结论对广泛的一类非连续核也会成立．

§2 证明细节

由于定理的证明很复杂，我们将证明的主要部分分解为一串引理．

引理 1 设 $0<h_n\to 0$．若存在函数 g 使(4)成立，则 F 在 R^1 上处处连续．

证明 证明很简单：若存在一点 x_0 使 $\triangle=F(x_0+)-F(x_0)>0$，则由大数定律有

$$N_n(x_0-h_n,\ x_0+h_n)/n\xrightarrow{P}\triangle$$

因而 $f_n(x_0)=N_n(x_0-h_n,\ x_0+h_n)/(2nh_n)\xrightarrow{P}\infty$，与 (4) 矛盾．

引理 2 设 $0<h_n\to 0$，$nh_n\to\infty$，且存在函数 g 使 (4) 成立，则 g 在 R^1 上一致连续．

证明 由引理 1 知 F 在 R^1 上连续，故不失普遍性可设样本 $X_1,\cdots,X_n$ 两两不同．现往证由 (1) 式定义的 f_n 有如下的性质：存在 $\eta>0$，致

$$|y_1-y_2|\leqslant\eta\Rightarrow|f_n(y_1)-f_n(y_2)|\leqslant 2/(nh_n) \tag{5}$$

事实上，若此结论不对，则对每个自然数 m 可找到两点 y_{m1} 和 y_{m2}，致

$$|y_{m1}-y_{m2}|\leqslant 1/m,\ 但\ |f_n(y_{m1})-f_n(y_{m2})|>2/(nh_n). \tag{6}$$

因为当 $|y|$ 很大时 $f_n(y)=0$，知 $\{y_{m1},\ y_{m2},\ m=1,2,\cdots\}$ 为有界集，必要时抽取子序列，不失普遍性可设

$$\lim_{m\to\infty} y_{m1}=\lim_{m\to\infty} y_{m2}=y_0$$

由 f_n 的定义易见它在 R^1 上左连续，又对任何 x，$\lim\limits_{y\downarrow x} f_n(y)$ 存在且等于 $f_n(x)$ 或 $f_n(x)\pm(2nh_n)^{-1}$．因而

$$\lim_{m\to\infty}\sup|f_n(y_{m1})-f_n(y_{m2})|$$

$$\leqslant \lim_{m\to\infty} \sup |f_n(y_{m1})-f_n(y_0)|+\lim_{m\to\infty} \sup |f_n(y_{m2})-f_n(y_0)|$$
$$\leqslant (2nh_n)^{-1}+(2nh_n)^{-1}=(nh_n)^{-1}$$

因而当m充分大时应有 $|f(y_{m1})-f(y_{m2})|<2/(nh_n)$，这与（6）矛盾，因而证明了满足(5)的$\eta$的存在。

现任给$\varepsilon>0$。因（4）成立，存在充分大的n，及样本$X_1,\cdots,X_n$。使由它作出的估计f_n满足条件

$$\sup_x |f_n(x)-g(x)|<\varepsilon/3$$

然后根据上述已证事实，找$\eta>0$致

$$|y_2-y_1|\leqslant\eta\Rightarrow|f_n(y_2)-f_n(y_1)|<2/(nh_n)<\varepsilon/3$$

这里用到 $\lim\limits_{n\to\infty} nh_n=\infty$。当 $|y_2-y_1|\leqslant\eta$ 时有

$$|g(y_2)-g(y_1)|\leqslant|g(y_2)-f_n(y_2)|+|f_n(y_2)-f_n(y_1)|+|f_n(y_1)-g(y_1)$$
$$<\varepsilon/3+\varepsilon/3+\varepsilon/3=\varepsilon$$

因而证明了本引理。

引理 3　设以下条件成立：

1. $0<h_n\to0$，$nh_n\to\infty$。
2. $h_n^*=(2r_n+1)h_n$，$n=1,2,\cdots$，r_n为非负整数，且$r_nh_n\to0$。
3. 存在函数g使（4）式成立。

以f_n^*记在f_n的定义（1）中，将h_n改为h_n^*所得的结果，则有

$$\sup_x|f_n^*(x)-g(x)|\to0\ a.e.,\ \text{当}\ n\to\infty \tag{7}$$

证明　由h_n^*和h_n的关系及f_n^*，f_n的定义，知对任何x有

$$f_n^*(x)=\frac{1}{2r_n+1}\sum_{i=-r_n}^{r_n} f_n(x+2ih_n)$$

由（4）得知当$n\to\infty$时

$$\sup_x\left|f_n^*(x)-\frac{1}{2r_n+1}\sum_{i=-r_n}^{r_n} g(x+2ih_n)\right|$$
$$\leqslant\frac{1}{2r_n+1}\sum_{i=-r_n}^{r_n}\sup_x|f_n(x+2ih_n)-g(x+2ih_n)|\to0,\ a.s. \tag{8}$$

由引理 2 知g在R^1上一致连续，又假定了$r_nh_n\to0$。故当$n\to\infty$时有

$$\sup_x\left|g(x)-\frac{1}{2r_n+1}\sum_{i=-r_n}^{r_n} g(x+2ih_n)\right|\to0 \tag{9}$$

由（8）、(9)即得(7)。引理证毕。

引理 4　设$0<h_n\to0$，$nh_n^2\to\infty$，且存在函数g使(4)成立，则g就是X的概率密度函数。

证明　由$h_n\to0$，$nh_n^2\to\infty$，知$nh_n\to\infty$，故由引理 2 知g在R^1上一致连续。记

$$F^*(x,h_n)=[F(x+h_n)-F(x-h_n)]/(2h_n) \tag{10}$$

则对任何 $\varepsilon>0$ 及固定的 x 有

$$\{|f_n(x)-F^*(x,h_n)|\geqslant\varepsilon\}\subset\{|N_n(x-h_n,\ x+h_n)/n-[F(x+h_n)-F(x-h_n)]|\geqslant 2h_n\varepsilon\}$$

注意到随机变量 $N_n(x-h_n,\ x+h_n)$服从二项分布 $B(n,p_n)$，其中

$$p_n=F(x+h_n)-F(x-h_n) \tag{11}$$

由 $nh_n^2\to\infty$ 得知当 $n\to\infty$ 时

$$P(|f_n(x)-F^*(x,h_n)|\geqslant\varepsilon)\leqslant p_n(1-p_n)/[n(2h_n\varepsilon)^2]\to 0$$

因而

$$f_n(x)-F^*(x,h_n)\xrightarrow{P}0,\ 当\ n\to\infty \tag{12}$$

由（12）与（4），得

$$\lim_{n\to\infty}F^*(x,h_n)=g(x),\ 对任何\ x\in R^1 \tag{13}$$

由引理1知 F 在 R^1 上连续，故可将 F 表为

$$F=F_{ac}+F_s \tag{14}$$

这里 F_{ac} 和F_s 分别是 F 的绝对连续部分和奇异部分．由(13)知

$$F'_{ac}(x)=g(x) \tag{15}$$

在 R^1 上对 Lebesgue 测度几乎处处成立．因此，有

$$F_{ac}(x)=F_{ac}(0)+\int_0^x g(t)dt$$

再由 g 在 R^1 上处处连续，即知(15)对任何 $x\in R^1$ 成立．因为 $f_n(x+h_n)-g(x+h_n)\xrightarrow{P}0$ 且 g 连续，有

$$f_n(x+h_n)\xrightarrow{P}g(x)\quad x\in R^1 \tag{16}$$

另一方面，与证明（12）一样的方法，可证

$$f_n(x+h_n)-F^*(x+h_n,h_n)\xrightarrow{P}0.\ x\in R^1 \tag{17}$$

由（16）、（17），得

$$\lim_{n\to\infty}F^*(x+h_n,h_n)=g(x),\ x\in R^1 \tag{18}$$

但

$$F^*(x+h_n,h_n)=\frac{1}{2h_n}[F_{ac}(x+2h_n)-F_{ac}(x)]+\frac{1}{2h_n}[F_s(x+2h_n)-F_s(x)] \tag{19}$$

由(15)、(18)、(19)，得

$$\lim_{n\to\infty}\frac{1}{2h_n}[F_s(x+2h_n)-F_s(x)]=0,\ x\in R^1$$

就是说，对每个 x，F_s 在 x 点有一个导出数为0．令 $\tilde{F}(x)=F_s(x)+x$，则 $\tilde{F}$ 连续，严增，且在每个点 x 有一个导出数为1．据实变函数论中周知的事实（见[6]，p.127，引理4），知对任何 $a<b$ 有 $\tilde{F}(b)-\tilde{F}(a)\leqslant b-a$，即 $F_s(b)-F_s(a)\leqslant 0$ 对任何 $a<b$．由于 F_s 非降，知 F_s 为一常数（可取为0），于是有 $F=F_{ac}$，即 F 为绝对连续，且由(15)知 $F'=g$，引理

证毕。

引理 5　设 $0<h_n\to 0$，但 $\liminf\limits_{n\to\infty} nh_n<\infty$。则不可能存在函数 g，使(4)成立。

证明　用反证法，设存在函数 g 使(4)成立。必要时取适当子序列，可设 $\lim\limits_{n\to\infty} nh_n=\lambda$存在有限。先设 $\lambda>0$。固定 x。仍定义 $F^*(x,h_n)$ 如(10)。必要时选取子序列，可设 $\lim\limits_{n\to\infty} F^*(x, h_n)=\mu\leqslant\infty$ 存在。分以下三种情况讨论：

1° $0<\mu<\infty$。

因为 $N_n(x-h_n,\ x+h_n)$ 服从二项分布 $B(n,p_n)$(p_n 见(11))，而 $\lim\limits_{n\to\infty} np_n=2\lim\limits_{n\to\infty} nh_n F^*(x,h_n)=2\lambda\mu$，$0<\lambda\mu<\infty$。以 Y 记服从 Poisson 分布 $\mathscr{P}(2\lambda\mu)$ 的随机变量，则当 $n\to\infty$ 时，$f_n(x)=N_n(x-h_n,\ x+h_n)/(nh_n)$ 的分布收敛于 Y/λ 的分布。这显然与 $f_n(x)\xrightarrow{P}g(x)$矛盾。

2° $\mu=\infty$

这时 $np_n\to\infty$，因而易证 $f_n(x)\xrightarrow{P}\infty$，也与 $f_n(x)\xrightarrow{P}g(x)$矛盾。

3° $\mu=0$

这时易见 $f_n(x)\xrightarrow{P}0$。这必须对一切 $x\in R^1$ 成立（否则存在一点 x，使上述情况 1° 或 2° 成立，而这已推出矛盾）。因此由(4)知 $g\equiv 0$，即

$$\sup_x|f_n(x)|\to 0,\ a.e.\ 当\ n\to\infty. \tag{20}$$

取 $a>0$ 充分大，致 $F(a)-F(-a)=\triangle>0$。对每个 n，找最小的奇自然数 $2r_n+1$，使 $(2r_n+1)h_n\geqslant a$。由 $h_n\to 0$，知

$$\lim_{n\to\infty}(2r_n+1)h_n=a. \tag{21}$$

定义随机变量

$$Z_n=\frac{1}{2r_n+1}\sum_{i=-r_n}^{r_n} f_n(2ih_n),$$

则由(20)知 $Z_n\to 0$，$a.e.$ 另一方面，易见

$$Z_n=\frac{1}{(2r_n+1)2h_n n}N_n(-(2r_n+1)h_n,\ (2r_n+1)h_n). \tag{22}$$

由 $0\leqslant N_n\leqslant n$ 及(21)，知$\{Z_n\}$一致有界，这与 $Z_n\to 0$ $a.e.$ 结合，得 $E(Z_n)\to 0$。但由(22)和(21)并注意 N_n 为二项分布，得

$$E(Z_n)=\frac{1}{(2r_n+1)2h_n}[F((2r_n+1)h_n)-F(-(2r_n+1)h_n)]$$

$$\geqslant\frac{1}{(2r_n+1)2h_n}\triangle\to\frac{\triangle}{2a}>0.$$

这与 $E(Z_n)\to 0$ 矛盾。

剩下考虑$\lim_{n\to\infty} nh_n=0$的情况．这个情况很明显：因为当$nh_n\to 0$时，对任给的M，当n充分大时，由(1)式定义的$f_n(x)$只能取0或大于M之值，且$f_n(x)$（对任何样本$X_1,\cdots,X_n$）不恒等于0．这样的f_n不可能一致收敛于任何函数g．这就完成了引理5的证明．

定理的证明 现在不难完成定理的证明．设$0<h_n\to 0$．由引理5知，有

$$\lim_{n\to\infty} nh_n=\infty \tag{23}$$

由（23）及（4），根据引理2，知g在R^1上一致连续．对每个n，取最小的奇自然数c_n，致$c_n^2h_n\geqslant 1$．由c_n的最小性，知

$$1\leqslant c_n^2h_n\leqslant 9 \tag{24}$$

令$h_n^*=c_nh_n$，$n=1,2,\cdots$．由（24）及$h_n\to 0$，知$h_n^*\to 0$．又由（23）知

$$nh_n^{*2}=nc_n^2h_n^2\geqslant nh_n\to\infty \tag{25}$$

定义

$$f_n^*(x)=N_n(x-h_n^*,\ x+h_n^*)/(2nh_n^*)$$

由h_n^*与h_n的关系，根据引理3，由(4)式可得

$$\sup_x|f_n^*(x)-g(x)|\to 0,\ a.e.\ 当\ n\to\infty \tag{26}$$

再由（25）及（26），根据引理4，即得$F'(x)=g(x)$在R^1上处处成立．定理证毕．

注 在定理中我们假定了$h_n\to 0$．我们不妨来看看，若此不成立，情况将如何，通过取子序列，我们可以限于考虑$\lim_{n\to\infty} h_n=h$存在的情况，又可分为$h=\infty$和$0<h<\infty$两款．

若$h=\infty$，则由f_n的定义立见，$f_n(x)$在$x\in R^1$上一致收敛于$g\equiv 0$．g当然不能是X的概率密度．

若$0<h<\infty$，我们可用Vapnik等在[7]中证明的一个结果的特例。它可写为：设$X_1,\cdots,X_n$为抽自分布F的独立随机样本，F_n为$X_1,\cdots,X_n$的经验分布，则对任给$\varepsilon>0$，有

$$P\{\sup_{a<b}|(F_n(b)-F_n(a))-(F(b)-F(a))|\geqslant\varepsilon\}\leqslant 16n^2\exp\left(-\frac{n\varepsilon^2}{8}\right)$$

由此可知，对任给$\varepsilon>0$有

$$\sum_{n=1}^{\infty}P\{\sup_{a<b}|(F_n(b)-F_n(a))-(F(b)-F(a))|\geqslant\varepsilon n^{-1/3}\}<\infty$$

因而当$n\to\infty$时

$$n^{1/3}\sup_{a<b}|(F_n(b)-F_n(a))-(F(b)-F(a))|\to 0,\ a.e.$$

因此以概率为1地，当n充分大时，对$x\in R^1$一致地有

$$|(F_n(x+h_n)-F_n(x-h_n))-(F(x+h_n)-F(x-h_n))|\leqslant n^{-1/3}$$

注意到$N_n(x-h_n,\ x+h_n)=n(F_n(x+h_n)-F_n(x-h_n))$，及$f_n$的定义，得知当$n$充分大时

对 x 一致地有

$$|f_n(x) - \frac{1}{2h_n}(F(x+h_n) - F(x-h_n))| \leqslant \frac{1}{2h_n} n^{-1/3} \tag{27}$$

为简单计，设 F 在 R^1 上处处连续（否则 F 肯定无密度）。则 F 在 R^1 一致连续，因而由 $\lim\limits_{n\to\infty} h_n = h > 0$ 可知

$$\lim_{n\to\infty} \frac{1}{2h_n}[F(x+h_n) - F(x-h_n)] = \frac{1}{2h}[F(x+h) - F(x-h)] \triangleq g(x) \tag{28}$$

在 R^1 上一致成立。由（27），（28），知这时（4）成立，且 g 定义如（28）。

即使 X 有概率密度 f，f 也不可能是 g。因为，由（28）知 g 在 R^1 上处处连续。若 X 有概率密度 $f=g$，则 $F'(x)$ 处处等于 $g(x)$。（28）可改写为

$$g(x) = \frac{1}{2h}\int_{x-h}^{x+h} g(t)dt, \quad \text{对任何 } x \in R^1 \tag{29}$$

但 g 为连续概率密度，且由（28）知 $\lim\limits_{|x|\to\infty} g(x) = 0$，知存在一点 a，使 $g(a) = \sup\limits_x g(x)$，且 $g(x) < g(a)$ 当 $x > a$。对 $x = a$，（29）式显然不能成立。这证明了 g 决不能是 X 的密度。

从实用的角度看，本注中讨论的情况，以及引理 5 中讨论的情况，都是学究式的。h_n 的有实用意义的取法，总要满足 $h_n \to 0$ 且 $nh_n \to \infty$。然而，引理 5 和本注中的数学论证使我们得以完全澄清这个关于 h_n 的选择问题。

参 考 文 献

[1] Rosenblatt, M., Remarks on some nonparametric estimates of a density function, *Ann. Math. Statist.* 27 (1956) pp. 832—837.

[2] Parzen, E., On estimation of a probability density function and mode, *Ann. Math. Statist.* 33 (1962) pp. 1065—1076.

[3] Nadaraya, E. A. On non-parametric estimates of density functions and regression corves, *Theor. Prob. Appl.* 10 (1965) pp. 186—190.

[4] Devroye, L. P. and Wagner, T. J., The Strong uniform consistency of kernel density estimates *Multivariate Analysis* V (1980) pp. 59—77.

[5] Schuster, E. F., Estimation of a probability density function and its derivatives, *Ann. Math. Statist.* 40 (1969) pp. 1187—1195.

[6] 复旦大学数学系，实变数函数论与泛函分析概要，1963 年。

[7] Vapnik, V. N. and Chervonenkis, A. Y. On the uniform convergence of the relative frequencies of events to their probabilities, *Theor. Prob. Appl.* 16 (1971) pp. 264—280.

Necessary Conditions for Uniform Convergence of Rosenblatt's Density Estimates

By *Chen Xiru* (陈希孺)

Abstract

Let $X_1, \cdots, X_n$ be iid. samples drawn from a population with probability density function f. The so-called kernel estimator of f has a form

$$f_n(x)=\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)\Big/(nh_n)$$

where $h_n>0$ is a chosen number, and K is a chosen function called the kernel. If

$$K(x)=\begin{cases}1/2, & \text{when } -1<x\leqslant 1,\\ 0, & \text{otherwise}\end{cases}$$

then the corresponding f_n is called Rosenblatt Estimator. For a class of continuous kernel functions and for h_n satisfying some condition, Schuster proved in [5] that the necessary condition for

$$\lim_{n\to\infty}\{\sup_x|f_n(x)-f(x)|\}=0, \quad a.e.$$

is that f is continuous uniformly on R^1. In this article we show that the same conclusion remains true for Rosenblatt Estimator, without any restriction on h_n.