ノンパラメトリック確率密度関数の再帰的推定

磯 貝 英 一

§0. はじめに

平均，分散が未知の正規分布のように，未知なパラメータを含み関数型が既知な分布から得られる標本に基づいてパラメータを推定する方法をパラメトリック法という．これに対して，関数型が未知な場合の推定方法をノンパラメトリック法という．このノンパラメトリック法は分布型が想定できない場合に有力となる．ところで，Rosenblatt [65] がカーネル法を用いてノンパラメトリック確率密度関数の推定問題を考えて以来，Parzen [61] を始め，非常に多くの研究者によってこの問題の研究が続けられてきた．これは，この問題が統計的推測において興味ある問題であり，また，パターン認識などの応用面においても重要だからである（[26]）．一方，推定方法としてカーネル法以外にもヒストグラム法，直交級数法などが提案され，研究がなされている（[27], [89] など）．

ところで，Rosenblatt や Parzen の推定量は非再帰的である．すなわち，標本が追加されたとき，推定量は最初から再計算しなければならない．また，再計算するためには，得られた標本をすべて保存しておかなければならない．この点を改良するために，Wolverton and Wagner [90] と Yamato [91] は再帰的推定量を提案した．この論説ではカーネル法を用いた再帰的推定量を中心にノンパラメトリック確率密度関数の推定問題について解説する．この推定問題に関しては，[17], [21], [57], [63], [74], [81] などの本にまとめられている．この他に，標本の大きさが確率的になる場合の逐次推定問題についても触れる．最後に，関連した問題としてノンパラメトリック回帰関数の推定問題について若干述べる．

§1. 問題と再帰的推定量

$\mathscr{B}_p$ を p 次元ボレル集合体とし，X と $\{X_n, n\geq 1\}$ はある確率空間 $(\Omega, \mathscr{F}, P)$ 上で定義された p 次元確率ベクトル列で，ともに $\mathscr{B}_p$ 上の共通した分布 λ_X を持つと仮定する．ここで，λ_X は R^p 上のルベーグ測度 μ に関して関数型が未知な確率密度関数（p. d. f. と略す）f を持ち，p 次元確率密度関数のある集合 $\mathcal{S}$ に属すると仮定する．確率密度関数の推定問題とは標本 $X^n=(X_1, \cdots, X_n)\in R^{pn}$ に基づいて $f(x)(x\in R^p)$ の推定量 $f_n(x)\equiv f_n(x;X^n)$ を構成することである．この $f_n(x)$ が各 $n(\geq 1)$ と X^n を固定した時，R^p 上の確率密度関数になるならば，$f_n(x)$ を確率密度推定量という．$\{f_n(x), n\geq 1\}$ が再帰的（recursive）であるとは，各 n に対して可測関数 Ψ_n が存在して $f_n(x)=\Psi_n(f_{n-1}(x), x, X_n)$ が成り立つことである．各 n と $j=1, \cdots, n$ に対して $g_{n,j};R^p\times R^p\to R$ を可測関数とする時，$f_n(x)=\sum_{j=1}^{n} g_{n,j}(x;X_j)$ で定義される推定量をカーネル（kernel）推定量という．

これまでに再帰的カーネル確率密度推定量としていくつかが提案され，その統計的性質が調べられている（[88]）．今，カーネルと呼ぶ実数値ボレル可測関数 $K(y)=K(y_1, \cdots, y_p)$ $(y=(y_1, \cdots, y_p)\in R^p)$ と bandwidth と呼ぶ正数列 $\{h_n, n\geq 1\}$ が任意に与えられているとする．以下に K と $\{h_n, n\geq 1\}$ を用いた f のカーネル推定量についてこれまでに提案されたものを列挙する．Rosenblatt [65] と Parzen [61] は1次元 ($p=1$) に対して，Cacoullos [7] は多次元 ($p\geq 1$) に対して次の非再帰的推定量を提案した．ここでは，これを Rosenblatt-Parzen estimator といい，略して単に RPE と呼ぶ．これ以後，他の推定量についても同様な略称を用いることにする．

(RPE) $$f_n(x)=\frac{1}{nh_n^p}\sum_{j=1}^{n}K\left(\frac{x-X_j}{h_n}\right).$$

さて，Wolverton and Wagner [90] と Yamato [91] は，非再帰的推定量 RPE を修正して，次の再帰的推定量 WWYE を提案した．

(WWYE) $$f_n(x)=\frac{1}{n}\sum_{j=1}^{n}K_j(x-X_j).$$

ここに，$K_j(y)=\frac{1}{h_j^p}K\left(\frac{y}{h_j}\right)$ である．WWYE は次のように再帰的に計算することができる．

$$f_n(x)=\left(1-\frac{1}{n}\right)f_{n-1}(x)+\frac{1}{n}K_n(x-X_n).$$

この他に，再帰的推定量として以下のような推定量 DE ([14]), BE ([4]), WDE ([87]), IE ([37]) が提案されている．

(DE) $$f_n(x)=\frac{1}{b_n}\sum_{j=1}^{n}H(h_j)K\left(\frac{x-X_j}{h_j}\right),$$

ここに，$b_n=\sum_{j=1}^{n}h_j^pH(h_j)$, $H:(0,\infty)\to(0,\infty)$ は任意の関数である．特に，$H(y)\equiv 1$ とした時の推定量を DE (1) と呼ぶ．

(BE) $$f_n(x)=\frac{1}{d_n}\sum_{j=1}^{n}\frac{h_j^p}{d_j}\sum_{i=1}^{j}K\left(\frac{x-X_j}{h_i}\right),$$

ただし，$d_n=\sum_{j=1}^{n}h_j^p$ である．

(WDE) $$f_n(x)=\frac{1}{nh_n^{p/2}}\sum_{j=1}^{n}\frac{1}{h_j^{p/2}}K\left(\frac{x-X_j}{h_j}\right),$$

(IE) $$f_0(x)\equiv K(x)$$
$$f_n(x)=(1-a_n)f_{n-1}(x)+a_nK_n(x-X_n)\quad(n\geq 1),$$

ただし，$\{a_n, n\geq 1\}$ は $a_n\leq 1$, $\lim_{n\to\infty}a_n=0$, $\sum_{n=1}^{\infty}a_n=\infty$ を満たす正数列である．推定量 IE は Watanabe [84] が与えたものを若干修正したものである．特に，$a_n=\frac{a}{n}$, $0<a\leq 1$ とおいた推定量 IE を IE (a) と呼ぶ時，推定量 IE (1) は推定量 WWYE になる．また，推定量 DE で $H(y)=y^{-p}$ とおくと，推定量 WWYE になる．もし K が確率密度関数であれば推定量 RPE, WWYE, DE, BE, IE は確率密度推定量になる．本論説ではカーネルを用いた再帰的推定量の統計的性質について解説する．推定量 RPE に関しては，その統計的性質について非常に多くの興味ある結果が得られているが，この論説では詳しく触れないことにする (例えば，[6], [18], [19], [66])．

最後に，この論説を通じて用いられる記号をまとめておく．$E(\cdot)$, $V(\cdot)$ は与えられた p.d.f. f に関する平均および分散を表す．'→ in prob.', '→ a.s.' はそれぞれ確率収束，概収束を表す．

$$\gamma_0=\gamma_1=1,\quad \gamma_n=\prod_{j=2}^{n}(1-a_j)\,(n\geq 2)$$

とおくと，推定量 IE は次のように書ける．

(1.1) $$f_n(x)=\gamma_n\sum_{j=0}^{n}a_j\gamma_j^{-1}K_j(x-X_j)\quad(n\geq 0).$$

ただし，$a_0=1,\ h_0=h_1,\ X_0=0,\ K_0(x)=K(x)$．$L_1=L_1(R^p, \mathcal{B}_p, \mu)$ は可積分な実数値関数のすべての集合を表し，$\|\cdot\|$ は R^p 上のユークリッドノルムを表す．本論説を通じて常にカーネル K は R^p 上で有界で可積分な実数値ボレル可測関数で，

$$\text{(K 1)} \qquad \|y\|^p|K(y)|\to 0 \ \ (\|y\|\to\infty) \quad \text{かつ} \quad \int K(y)dy=1$$

を満たし，積分は全範囲 R^p でとられるとし，積分範囲を省略する．また，K の動径的優関数 (radial majorant) φ を $\varphi(x)=\sup_{\|y\|\geq\|x\|}|K(y)|$ で定義し，bandwidth $\{h_n, n\geq 1\}$ は 0 に収束する正数列とする．

§2. 良さの規準

これまでに，いろいろな推定量の良さを比較するための基準が考えられている ([86])．ここではそのうちのいくつかを紹介する．以下では，$f_n(x)\equiv f_n(x;X^n)$ を与えられた推定量とし，$n\to\infty$ の時の収束について考える．任意の点 x と任意の $f\in\mathcal{S}$ に対して $f_n(x)$ の平均 2 乗誤差 (mean squared error, 略して MSE) $\mathrm{MSE}(f_n(x))=E(f_n(x)-f(x))^2$ が 0 に収束する時，推定量 f_n は MSE の意味で各点一致性を持つという．この収束が x に関して一様ならば，f_n は MSE の意味で一様一致性を持つという．任意の点 x と任意の $f\in\mathcal{S}$ に対して $f_n(x)$ が $f(x)$ に確率 (概) 収束する時，f_n は各点弱 (各点強) 一致性を持つという．この収束が x に関して一様ならば，f_n は一様弱 (一様強) 一致性を持つという．ところで，MSE の意味で各点一致性を持てば各点弱一致性を持つ．また，各点弱一致性よりも各点強一致性が望ましく，各点一致性よりも一様一致性が望ましい．正数 r に対して f_n が L_r の意味で弱 (強) 一致性を持つとは，任意の $f\in\mathcal{S}$ に対して $\int|f_n(x)-f(x)|^r dx$ が 0 に確率 (概) 収束することである．特に，平均積分 2 乗誤差 (mean integrated squared error, 略して MISE) $E\left[\int|f_n(x)-f(x)|^2dx\right]$ が 0 に収束する時，f_n は MISE の意味で一致性を持つという．

§3. 独立な観測

3.1. 漸近不偏性

第 3 節では，p 次元確率ベクトル $X_1,\cdots,X_n$ は独立で同一分布に従い，p.d.f. f をもつと仮定する．ここで，すべての x とすべての f に対して，$Ef_n(x)=f(x)$ となる f の確率密度推定量 f_n が存在しないことが知られている ([65], [72])．従って，与えられた推定量が漸近不偏性をもつかどうか，すなわち，$\lim_{n\to\infty}Ef_n(x)=f(x)$ が成り立つかどうかが興味の対象の 1 つになる．次の補題は漸近的性質を示すために基本となる ([7])．

補題 3.1. $g\in L_1$ と正数 h に対して $K_h(y)=\dfrac{1}{h^p}K\left(\dfrac{y}{h}\right)$,

$$g*K_h(x)=\int g(x-y)K_h(y)dy=\int K_h(x-y)g(y)dy$$

($*$ はたたみこみ) とおくと，g の任意の連続点 x に対して $\lim_{h\to+0}g*K_h(x)=g(x)\int K(y)dy$ が成り立つ．

なお，上の補題では，$\int K(y)dy=1$ は必ずしも満たされなくてもよい．Stein ([76], p. 62～63)

により g の連続性を仮定しない場合には次の結果が得られる．

補題 3.2. $g\in L_1$ で，K は動径的優関数 $\varphi\in L_1$ を持つとする．この時，ほとんどすべての x に対して $\lim_{h\to+0} g*K_h(x)=g(x)$ が成り立つ．

ここで，$\varphi\in L_1$ であるための1つの十分条件は，§1 で仮定した K が，ある $\varepsilon>0$ に対して，$K(x)=O(\|x\|^{-p-\varepsilon})\,(\|x\|\to\infty)$ を満たすことである．

さて，漸近不偏性について考える．補題 3.1 と Toeplitz の補題（[52] p. 250）を用いると，漸近不偏性に関する結果が得られる．

定理 3.1. x を任意の f の任意の連続点とする．この時，推定量 RPE と IE は漸近不偏性を持つ．もし $\lim_{n\to\infty} b_n=\infty$ が満たされれば，推定量 DE は漸近不偏性を持つ．また，もし $\lim_{n\to\infty} d_n=\infty$ が満たされれば，推定量 BE は漸近不偏性を持つ．

3.2. 一 致 性

$\mathrm{MSE}(f_n(x))=V(f_n(x))+B_n^2(x)$ $(B_n(x)=Ef_n(x)-f(x))$ が成り立つので補題 3.1 と定理 3.1 を用いると MSE の意味での各点一致性に関する次の結果を得る．

定理 3.2. x を任意の f の任意の連続点とする．この時，もし $\lim_{n\to\infty} nh_n^p=\infty$ が満たされれば，推定量 RPE は MSE の意味で各点一致性を持つ．もし $\lim_{n\to\infty} b_n=\infty$ と $\lim_{n\to\infty} b_n^{-2}\sum_{j=1}^{n} h_j^p H^2(h_j)=0$ が満たされれば，推定量 DE は MSE の意味で各点一致性を持つ．また，もし $\lim_{n\to\infty} a_n h_n^{-p}=0$ が満たされれば，推定量 IE は MSE の意味で各点一致性を持つ．

関係式 $f_n(x)-f(x)=(f_n(x)-Ef_n(x))+B_n(x)$ と定理 3.1 を利用すると各点強一致性に関する次の結果が得られる（[37], [63]）．

定理 3.3. x を任意の f の任意の連続点とする．この時，もしすべての定数 $c>0$ に対して $\sum_{n=1}^{\infty}\exp(-cnh_n^p)<\infty$ が満たされれば，推定量 RPE は各点強一致性と MSE の意味での各点一致性を持つ．また，もし $\sum_{n=1}^{\infty} a_n^2 h_n^{-p}<\infty$ が満たされれば，推定量 IE は各点強一致性と MSE の意味での各点一致性を持つ．

f の連続性を仮定しない場合には補題 3.2 を用いると次の結果が得られる（[21]）．

定理 3.4. f を任意の p. d. f. とし，K は R^p 上で非負で，動径的優関数 $\varphi\in L_1$ を持つとする．この時，$\lim_{n\to\infty} nh_n^p=\infty$ ならば，ほとんどすべての x に対して推定量 WWYE は各点弱一致性を持つ．また，もし $\lim_{n\to\infty}\dfrac{nh_n^p}{\log_2 n}=\infty$ ならば，ほとんどすべての x に対して推定量 WWYE は各点強一致性を持つ．ここで，$\log_2 n=\log\log n$ である．

カーネルのフーリエ変換を用いて推定量 IE の一様強一致性を示すことができる（[37]）．今，

$$k(t)=\int\exp\Bigl(i\sum_{j=1}^{p} t_j y_j\Bigr)K(y)dy,\quad i^2=-1,\quad A_n=\gamma_n^2\sum_{j=1}^{n} a_j^2\gamma_j^{-2}$$

とおく．

定理 3.5. f は R^p 上で一様連続な任意の p. d. f. であると仮定する．K は R^p 上で連続で，$|k(t)|\in L_1$ はすべての $u\neq0\in R^p$ に対し $R(u)=\{t=qu; q>0\}$ 上で非増加であるとする，すなわち，$\|t_1\|\leq\|t_2\|$ なる $t_1, t_2\in R(u)$ に対して $|k(t_1)|\geq|k(t_2)|$ とする．また，h_n は単調減少で，

$$\lim_{n\to\infty}\frac{h_n}{h_{n+1}}=1,\quad \sum_{n=1}^{\infty}(a_n h_n^{-p})^2<\infty,\quad \sum_{n=1}^{\infty}\frac{A_n}{h_n^{2p-1}}\left|\frac{1}{h_{n+1}}-\frac{1}{h_n}\right|<\infty$$

を満たすとする．このとき，推定量 IE は一様強一致性かつ MSE の意味で一様一致性をもつ．

注意 3.1.　1次元の場合，K と h_n に関するある条件の下では，f が一様連続であることが推定量 RPE が一様強一致性を持つための必要十分条件である（[71]）．

$I(A)$ を A の定義関数とするとき，L_1 の意味での一致性に関する結果が次で得られる（[21]）．

定理 3.6.　K は R^p 上で非負で，動径的優関数 $\varphi \in L_1$ を持ち，$\{h_n, n \geq 1\}$ を正数列とする時，次は同値になる．

（i）　すべての f とほとんどすべての x に対して DE(1) は各点強一致性をもつ．

（ii）　ある f とほとんどすべての x に対して DE(1) は各点弱一致性をもつ．

（iii）　$\lim_{n\to\infty} b_n = \infty$ かつ任意の $\varepsilon > 0$ に対して $\lim_{n\to\infty} \sum_{j=1}^{n} h_j^p I(h_j > \varepsilon)/b_n = 0$ が成り立つ．

（iv）　すべての f に対して DE(1) は L_1 の意味で強一致性を持つ．

（v）　ある f に対して DE(1) は L_1 の意味で弱一致性を持つ．

推定量 WWYE に対する L_1 の意味での一致性に関しても論じられている（[16]）．

3.3.　漸近分散

漸近分散の観点から推定量の比較を行う．

$$I_0 = f(x) \int K^2(y)\,dy \tag{3.1}$$

とおく時，関係式

$$V\left(K\left(\frac{x-X}{h}\right)\right) = h^p f * (K^2)_h(x) - h^{2p}(f * K_h(x))^2$$

と補題 3.1 を用いると漸近分散に関する次の結果が得られる（[4], [35] など）．

定理 3.7.　x を任意の f の任意の連続点とする．

（i）　推定量 RPE, WDE に対して

$$\lim_{n\to\infty} nh_n^p V(f_n(x)) = I_0$$

が成り立つ．

（ii）　$\lim_{n\to\infty} \frac{nh_n^p}{b_n^2} = 0,\ \lim_{n\to\infty} \frac{nh_n^p}{b_n^2} \sum_{j=1}^{n} h_j^p H^2(h_j) = \alpha$
を満たす正数 α が存在すれば，推定量 DE に対して

$$\lim_{n\to\infty} nh_n^p V(f_n(x)) = \alpha I_0$$

が成り立つ．

（iii）　h_n は単調減少で，ある $a \in \left(\frac{1}{2}, 1\right]$ に対して

$$\lim_{n\to\infty} n^{1-2a} h_n^p \sum_{j=1}^{n} j^{2(a-1)} h_j^{-p} = \beta$$

を満たす正数 β が存在すれば，推定量 IE(a) に対して

$$\lim_{n\to\infty} nh_n^p V(f_n(x)) = a^2 \beta I_0$$

が成り立つ．

注意 3.2.　$H(y) \equiv 1$ のとき，$\frac{nh_n^p}{b_n^2} \sum_{j=1}^{n} h_j^p H^2(h_j)$ は最小になるから α も最小になる．

さて，推定量を比較するために記号を導入する．f_n, g_n を2つの推定量とする時，$f_n <_1 g_n$

($f_n <_2 g_n$) であるとは，ある定数 $\eta\in(0,1)$ と正の整数 N が存在して，すべての $n\geq N$ に対して $V(f_n)\leq\eta V(g_n)$ ($\mathrm{MSE}(f_n)\leq\eta\,\mathrm{MSE}(g_n)$) が成り立つことである．従って，$f_n<_1 g_n$($f_n<_2 g_n$) の時，漸近分散（漸近 MSE）の観点から f_n は g_n よりも良い推定量であるといえる．ここで，$h_n=n^{-r/p}\left(0<r<\dfrac{1}{2}\right)$ の場合に漸近分散の観点から推定量を比較する．この場合，$H(y)\equiv1$ ならば $\alpha=1-r$, $\beta=(2a+r-1)^{-1}$ として定理 3.7 の条件は満たされる．従って，$f(x)>0$ を満たす f の連続点 x に関しては

$$\mathrm{DE}(1),\ \mathrm{IE}(1-r)<_1\mathrm{WWYE}<_1\mathrm{WDE},\ \mathrm{RPE}$$

が成り立つ．

3.4. MSE の漸近展開

この節では1次元 ($p=1$) の場合について推定量 RPE, DE(1), IE(a) の MSE の漸近展開を与え，漸近 MSE の観点から推定量の比較をする．このような議論は，例えば，[15] でなされている．以下では K は

$$\text{(K2)}\qquad K(y)=K(-y)\quad(y\in R),\quad \int y^2|K(y)|dy<\infty$$

を満たし，f は R 上で2回連続微分可能で，f'' は有界であると仮定する．この時，テイラーの定理を用いると次が成り立つ．

$$f*K_h(x)=f(x)+I_1h^2+o(h^2)\quad(h\to0),$$
$$f*(K^2)_h(x)=I_0+I_2h^2+o(h^2)\quad(h\to0).$$

ただし，$I_m=\dfrac{f''(x)}{2}\displaystyle\int y^2K^m(y)dy$ ($m=1,2$), I_0 は (3.1) で与えられたものである．この関係式から推定量 $f_n(x)$ の bias $B_n(x)$ の漸近展開に関する次の結果が得られる．

定理 3.8. f_n を推定量 RPE, DE(1), IE(a) の1つとする時，$B_n(x)=I_1\xi_n+o(\xi_n)$ $(n\to\infty)$ が成り立つ．ただし，

(i) RPE に対して $\xi_n=h_n^2$,

(ii) DE(1) に対して $\sum_{n=1}^{\infty}h_n^3=\infty$ ならば $\xi_n=b_n^{-1}\sum_{j=1}^{n}h_j^3$,

(iii) IE(a) に対して $\sum_{n=1}^{\infty}n^{a-1}h_n^2=\infty$ ならば

$$\xi_n=an^{-a}\sum_{j=1}^{n}j^{a-1}h_j^2.$$

ここで，$h_n=n^{-r}\left(0<r<\dfrac{1}{3}\right)$ の場合を考えると，定理 3.7 と定理 3.8 より MSE の漸近展開が求まる．

定理 3.9. $c_n=o(n^{r-1}+n^{-4r})$ とおくと，次が成り立つ．

$$\mathrm{MSE(RPE)}=I_0n^{r-1}+I_1^2n^{-4r}+c_n,$$

$$\mathrm{MSE(DE(1))}=(1-r)I_0n^{r-1}+\left(\frac{1-r}{1-3r}\right)^2I_1^2n^{-4r}+c_n,$$

$$\mathrm{MSE(WWYE)}=\frac{I_0}{1+r}n^{r-1}+\frac{I_1^2}{(1-2r)^2}n^{-4r}+c_n,$$

$$\mathrm{MSE}(\mathrm{IE}(1-r))=(1-r)I_0 n^{r-1}+\left(\frac{1-r}{1-3r}\right)^2 I_1^2 n^{-4r}+c_n.$$

注意 3.3. (i) 漸近的に各 MSE の右辺を最小にする r は $r=\frac{1}{5}$ である．さらに，漸近的に最適な bandwidth の選び方も考えられている ([78] など)．(ii) 漸近的に最適なカーネルの選び方についてもいろいろと議論されている ([11] など)．

さて，定理 3.9 を用いて漸近 MSE の観点から推定量を比較する．$0<r<\frac{1}{5}$ の時，$f''(x)\neq 0$ を満たす x に関しては，$\mathrm{RPE}<_2\mathrm{WWYE}<_2\mathrm{DE}(1), \mathrm{IE}(1-r)$ が成り立つ．$\frac{1}{5}<r<\frac{1}{3}$ の時，$f(x)>0$ を満たす x に関しては，$\mathrm{DE}(1)$, $\mathrm{IE}(1-r)<_2\mathrm{WWYE}<_2\mathrm{RPE}$ が成り立つ．$r=\frac{1}{5}$ の時はこの h_n に対しては必ずしも比較はできない．しかし，$h_n=\left(\frac{I_0}{4I_1^2}\right)^{1/5}n^{-1/5}$ と 3.5 節の (3.2) で与えられるカーネル K_0 を用いると，$f(x)>0$ かつ $f''(x)\neq 0$ を満たす x に関しては，$\mathrm{RPE}<_2\mathrm{WWYE}<_2\mathrm{DE}(1)$ が成り立つ ([88])．

なお，MISE の観点からカーネル法とヒストグラム法の MISE は次のようになる．f は R 上で 2 回連続微分可能で，f'' は有界とし，$f, f', f''\in L_2$ を仮定する．K は p.d.f. で $K(y)=K(-y)$ $(y\in R)$, $\int y^2K(y)dy=1$ を満たすとし，f_n を推定量 RPE とする．この時，$L=\int K^2(y)dy$, $M=\int(f''(y))^2dy$ とおいて，漸近的に最適な $h_n=\left(\frac{L}{M}\right)^{1/5}n^{-1/5}$ を考えると，$\mathrm{MISE}(f_n)=\left\{\frac{5}{4}M^{1/5}L^{4/5}\right\}n^{-4/5}+o(n^{-4/5})$ が成り立つ．この右辺を最小にするカーネルは (3.2) で与えられる ([23])．一方，$Q=\{C_j\}$ を R の分割 ($\bigcup_j C_j=R$; すべての $i\neq j$ に対して $C_i\cap C_j=\emptyset$) とし，$\mathcal{P}$ を R の分割のすべての集合とする．今，ある分割 $Q=\{C_j\}$ に対して，標本 $X_1,\cdots,X_n$ に基づく $x\in C_j$ におけるヒストグラムを $H_n(x|Q)=n^{-1}\sum_{i=1}^n I\{X_i\in C_j\}/|C_j|$ で定義する．ただし，$|C_j|$ は級区間 C_j の長さを表す．この時，このヒストグラム $H_n(x|Q)$ を用いた時の MISE を $\mathrm{MISE}(n,Q)$ とすると，

$$\inf_{Q\in\mathcal{P}}\mathrm{MISE}(n,Q)=\left\{\left(\frac{3}{2}\right)\left(\frac{1}{6}\right)^{1/3}\int|f'(y)f(y)|^{2/3}dy\right\}n^{-2/3}+o(n^{-2/3})$$

が成り立つ ([47], [48])．

3.5. 漸近正規性

簡単のために $h_n=n^{-r/p}\left(\frac{p}{p+2}<r<1,\ 2a+r>1\right)$ に対して推定量 $\mathrm{IE}(a)$ の漸近正規性を考える．(一般の場合は [39] を参照)．カーネル K は

(K3)
$$\int\|y\||K(y)|dy<\infty$$

を満たすとする．

定理 3.10. f は R^p 上で 1 回偏微分可能で，第 1 階偏導関数 $\frac{\partial f(x)}{\partial x_j}$ $(j=1,\cdots,p)$ は有界であると仮定する．$q\geq 1$ は任意に与えられた整数で，$x_1,\cdots,x_q$ は $f(x_j)>0$ $(j=1,\cdots,q)$ を満たす任意の異なる点とする．この時，$(nh_n^p)^{1/2}(f_n(x_1)-f(x_1),\cdots,f_n(x_q)-f(x_q))$ は平均ベクトル 0，共分散行列 Γ の q 次元正規分布に法則収束する．ただし

$$\Gamma=\frac{a^2}{2a+r-1}\int K^2(y)dy\begin{pmatrix}f(x_1) & & 0\\ & \ddots & \\ 0 & & f(x_q)\end{pmatrix}.$$

注意 3.4. $(nh_n^p)^{1/2}(f_n(x_j)-f(x_j))$ $(j=1,\cdots,q)$ は漸近的に独立に平均0，分散 $\frac{a^2}{2a+r-1}\cdot f(x_j)\int K^2(y)dy$ の正規分布に従うことがわかる．

カーネルの具体的な例をあげよう．以下の $K_0(t)$ $(t\in R)$ に対して，$K(y)=\prod_{j=1}^{p}K_0(y_j)$ とおくと，K は (K1), (K2), (K3) を満たす．特に，(3.3) は定理 3.5 の条件をも満たす．

$$K_0(t)=\frac{1}{2}I(|t|\le 1),\quad K_0(t)=(1-|t|)I(|t|\le 1),$$

$$K_0(t)=\frac{3}{4\sqrt{5}}\left(1-\frac{t^2}{5}\right)I(|t|\le\sqrt{5}), \tag{3.2}$$

$$K_0(t)=(2\pi)^{-1/2}\exp\left(-\frac{t^2}{2}\right),\quad K_0(t)=\frac{1}{2}\exp(-|t|). \tag{3.3}$$

§4. 従属な観測

$\{X_t;t=0,\pm1,\pm2,\cdots\}$ は R^p に値をとる強定常過程で，p.d.f. f を持つと仮定する．$\mathcal{M}_a^b(a\le b)$ は $X_a,\cdots,X_b$ から生成される σ-加法族を表し，$L_2(\mathcal{M}_a^b)$ は $E(U^2)<\infty$ で $\mathcal{M}_a^b$-可測な確率変数 U のすべての集合を表す．ここで，従属性を表す条件としていくつかの混合条件を定義する．

定義 4.1. (i) $\{X_t\}$ が強混合条件を満たすとは

$$\alpha(n)=\sup_{A\in\mathcal{M}_{-\infty}^0,B\in\mathcal{M}_n^\infty}|P(A\cap B)-P(A)P(B)|\to 0\quad(n\to\infty)$$

が成り立つことである．

(ii) $\{X_t\}$ が絶対正則条件を満たすとは

$$\beta(n)=E\{\sup_{A\in\mathcal{M}_n^\infty}|P(A|\mathcal{M}_{-\infty}^0)-P(A)|\}\to 0\quad(n\to\infty)$$

が成り立つことである．ここに，$P(\cdot|\mathcal{M}_{-\infty}^0)$ は条件付き確率を表す．

(iii) $\{X_t\}$ が完全正則条件を満たすとは

$$\gamma(n)=\sup\frac{|E(\xi\eta)-E(\xi)E(\eta)|}{(V(\xi)V(\eta))^{1/2}}\to 0\quad(n\to\infty)$$

が成り立つことである．ここに，上限はすべての $\xi\in L_2(\mathcal{M}_{-\infty}^0)$ と $\eta\in L_2(\mathcal{M}_n^\infty)$ についてとる．この時，絶対正則条件あるいは完全正則条件を満たせば強混合条件を満たす．これらの混合条件以外にもいろいろな混合条件が定義され，その性質が調べられている（例えば，[92]）．さて，Györfi ([32]) は推定量 WWYE に対し，L_2 の意味での強一致性を示した．Masry ([54]) は WWYE, WDE に対し，強混合条件と完全正則条件の下で MSE の意味での各点一致性および漸近正規性を与えた．また，Takahata ([80]) は RPE, WWYE に対し，絶対正則条件の下で一様強一致性に関する収束の速さを論じた．多次元の場合には推定量 WWYE $f_n(x)$ に対して，Masry ([53]) は強混合条件の下で各点強一致性および収束の速さに関する次の結果を与えた．

定理 4.1. K は動径的優関数 $\varphi\in L_1$ を持つとする．

(i) ある $\delta>0$ とある $r>2$ に対して

$$\sum_{n=1}^{\infty}(\log n)(\log_2 n)^{1+\delta}[\alpha(n)]^{1-(2/r)}\sum_{j=n}^{\infty}\frac{1}{j^2h_j^{2p(1-1/r)}}<\infty$$

かつ $\sum_{n=1}^{\infty}\frac{1}{n^2h_n^{2p(1-1/r)}}<\infty$ が満たされれば，ほとんどすべての $x\in R^p$ に対して WWYE は各点強

一致性を持つ.

（ii） f は R^p 上で2回偏微分可能で，すべての第2階偏導関数は有界かつ連続であるとする. K は，さらに

$$\int y_j K(y)dy=0 \ (j=1,\cdots,p),\quad \int \|y\|^2|K(y)|dy<\infty$$

を満たすとし，ある $r>2$ に対して

$$\sum_{n=1}^{\infty}(\log n)(\log_2 n)[\alpha(n)]^{1-(2/r)}<\infty$$

を仮定する．この時，$h_n=n^{-1/\gamma}(\gamma=4+2p(1-1/r))$ とすると，すべての x とすべての $\delta>0$ に対して

$$\left\{\frac{n^{4/\gamma}}{(\log n)(\log_2 n)^{1+\delta}}\right\}^{1/2}(f_n(x)-f(x))\to 0 \text{ a.s. } (n\to\infty)$$

が成り立つ.

Masry and Györfi ([55]) は完全正則条件の下で WWYE に対して各点強一致性に関する収束の速さを論じた．Tran ([82]) は1次元の場合に [53] の結果を改良した．さらに，多次元の場合に WWYE に対し，Tran ([83]) は別の従属条件の下で一様強一致性に関する収束の速さおよび漸近正規性を示した．この他，マルコフ過程に関しても [5], [43], [58], [59], [68] などで議論されている.

§5. 逐次推定

$X_1, X_2, \cdots$ はある確率空間 $(\Omega, \mathcal{F}, P)$ 上で定義された独立で同一分布に従う p 次元確率ベクトル列で，p.d.f. f を持つと仮定する．$\{\mathcal{F}_n, n\geq 1\}$ を $\mathcal{F}$ の部分 σ-加法族の増大列とする．$\{1, 2, \cdots, \infty\}$ の値をとる確率変数 N が $\{N\leq n\}=\{\omega\in\Omega; N(\omega)\leq n\}\in\mathcal{F}_n$, $n=1,2,\cdots$ を満たす時，($\mathcal{F}_n$−) 停止時間であるという．ところで，標本の大きさが確率的になる場合が実際しばしば起こる．たとえば，t 時間内に店に来る客を観測してある量を推定したい時，客の数は確率変数とみなされ，従って，標本の大きさは確率的になる．本節ではこのように標本の大きさが確率変数になる場合の確率密度関数の推定を扱う．以下では $\mathcal{F}_n$ を $(X_1,\cdots,X_n)$ から生成される σ-加法族とする．N を1つの停止時間，$f_n(x)$ を固定した大きさ n の標本に基づく $f(x)$ の推定量とする時，逐次推定量 $f_N(x)$ を，

$$f_N(x)=\begin{cases} f_n(x), & N=n \text{ の時},\\ \infty, & N=\infty \text{ の時}\end{cases}$$

と定義する．逐次推定量 $f_N(x)$ を扱ったのは Srivastava [75] が最初のように思われる．Davies and Wegman [13] は具体的に停止時間を与え，逐次推定量と停止時間のモーメントについて論じた．Carroll [8] は Chow and Robbins [10], Farrell [25], Sen and Ghosh [73] の考え方に基づいて，幅が $2d$ の $f(x)$ の信頼区間に関するタイプⅠ，タイプⅡの2種類の停止時間を導入した．Wegman and Davies [87] は別の停止時間を，Stute [79] は推定量 RPE を用いてタイプⅡの停止時間を，Isogai [36] は逐次推定量 IE(a) を用いてタイプⅠの停止時間を考えた．Koronacki and Wertz [49] は推定量 WWYE を用いて積分2乗誤差 (integrated squared error) がある与えられた値以下になる確率を $1-\alpha$ 以上にするための停止時間を導入した.

この節では1次元の場合についてタイプIの停止時間を用いた $f(x)$ の逐次推定問題を考える([36]). 以下では K は非負で (K2) を満たし, h_n は次の条件を満たすとする.

$$h_n\downarrow 0,\ nh_n\uparrow\infty,\ \frac{h_n}{h_{n+1}}\to 1(n\to\infty),\ \sum_{n=1}^{\infty}(n^2h_n)^{-1}<\infty,$$

$$\lim_{n\to\infty}\frac{\log n}{(nh_n)^{1/2}}=0,\ \lim_{n\to\infty}nh_n^3=0,\ \sum_{j=1}^{n}j^{a-1}h_j=o(n^ah_n),$$

与えられた $a\in(0,1]$ に対して

$$\lim_{n\to\infty}n^{1-2a}h_n\sum_{j=1}^{n}j^{2(a-1)}h_j^{-1}=\beta$$

を満たす $\beta\in(0,\infty)$ が存在する.

今, $\alpha\in(0,1)$ と $d>0$ が任意に与えられたとする. ここで Φ を標準正規分布の分布関数とし,

$$B=a^2\beta\int K^2(y)dy,\ b=B^{1/2}\Phi^{-1}\left(1-\frac{\alpha}{2}\right) \tag{5.1}$$

とおく. さて, $f_n(x)$ を再帰的推定量 IE(a) $\left(a\in\left[\frac{2}{3},1\right]\right)$ とし, x は $f(x)>0$ を満たすとする. この $f_n(x)$ を用いて十分小さい $d>0$ に対し, 漸近的に $1-\alpha$ の信頼係数を持ち, 幅が $2d$ の $f(x)$ の信頼区間を与えるために, 停止時間 $N_1(d)=N_1(d,x,\mathrm{IE}(a))$ を次のように定義する. もし

$$nh_n(d/b)^2\geq f_n(x)>0 \tag{5.2}$$

を満たす最小の正整数 n が存在すれば $N_1(d)=n$ とし, そのような n が存在しなければ $N_1(d)=\infty$ とする. $n(d)$ は $nh_n(d/b)^2\geq f(x)$ を満たす最小の正整数 n とする. そこで, 幅が $2d$ の $f(x)$ の信頼区間として, $I_{N_1(d)}(x)=[f_{N_1(d)}(x)-d, f_{N_1(d)}(x)+d]$ を考える. まず, 停止時間 $N_1(d)$ に関する結果として次が得られる.

定理 5.1. x を任意の f の任意の連続点とする. この時, すべての $d>0$ に対して $P\{N_1(d)<\infty\}=1$ かつ

$$\lim_{d\to+0}N_1(d)=\infty\ \text{a.s.},\ \lim_{d\to+0}\frac{N_1(d)h_{N_1(d)}d^2}{b^2f(x)}=1\ \text{a.s.},$$

$$\lim_{d\to+0}\frac{N_1(d)h_{N_1(d)}}{n(d)h_{n(d)}}=1\ \text{a.s.} \tag{5.3}$$

が成り立つ. また, $f_{N_1(d)}$ は各点強一致性を持つ.

次に, (1.1), 定理 3.10, 添数が確率変数の場合の極限定理に関する結果 ([3]) を用いると, 逐次推定量 $f_{N_1(d)}(x)$ の漸近正規性が次で得られる.

定理 5.2. f は R 上で1回微分可能で, f' は有界であると仮定する. また,

$$\frac{N_1(d)}{n(d)}\to 1\quad(d\to+0)\ \text{in prob.} \tag{5.4}$$

を仮定する. この時, 任意の x に対して $d\to+0$ として $(N_1(d)h_{N_1(d)})^{1/2}(f_{N_1(d)}(x)-f(x))$ は平均0, 分散 $Bf(x)$ の正規分布に法則収束する.

(5.4) の条件をより一般化した場合の漸近正規性についても議論されている ([46]). 定理 5.2 を用いると信頼区間 $I_{N_1(d)}(x)$ は Chow and Robbins [10] の意味で漸近的一致性を持つことが次でわかる.

系 5.1. 定理5.2の条件の下で

(5.5)
$$\lim_{d\to+0} P\{f(x)\in I_{N_1(d)}(x)\}=1-\alpha$$
が成り立つ．

ここで

(5.6)
$$h_n=n^{-r},\ \frac{1}{3}<r<a$$
を考える．(5.3) より (5.4) が成り立つから，次の結果が得られる．

系 5.2. f は R 上で1回微分可能で，f' は有界であると仮定する．この時，(5.6) の下で任意の x に対して $d\to+0$ として $\dfrac{(1-r)bf(x)}{B^{1/2}dg(d)}(N_1(d)-g(d))$ は標準正規分布に法則収束する．ただし，$g(d)=(b^2f(x)/d^2)^{1/(1-r)}$ である．

次に，(5.5) における収束の速さを考えよう．f_n として推定量 WWYE を，さらに $h_n=n^{-r}$, $\frac{1}{5}<r<1$ を考える．(5.2) の条件 $f_n(x)>0$ を若干修正して，停止時間 $N_2(d)=N_2(d,x,$ WWYE) を次で定義する．もし
$$n^{1-r}(d/b)^2\geq f_n(x)+\frac{1}{n}$$
を満たす最小の正整数 n が存在すれば $N_2(d)=n$ とし，そのような n が存在しなければ $N_2(d)=\infty$ とする．添数が確率変数の場合の中心極限定理の収束の速さに関する結果（[69]）を用いると次の結果が得られる（[41]）．

定理 5.3. f は R 上で2回微分可能で，f'' は有界であると仮定する．この時，任意の x と任意の $d>0$ に対して $P\{N_2(d)<\infty\}=1$ かつ

(5.7)
$$P\{f(x)\in I_{N_2(d)}(x)\}=1-\alpha+o(d^\eta)\quad(d\to+0)$$
が成り立つ．ここに

(5.8)
$$\eta=\min\left\{\frac{r}{2},\ \frac{2(1-r)}{5(2-r)},\ \frac{5r-1}{1-r}\right\}$$
である．

注意 5.1. (5.8) の η は $r=r_0=(7-\sqrt{29})/5$ で最大値 $(7-\sqrt{29})/10$ をとる．注意 3.3 では $h_n=n^{-r}$ に対する漸近的に最適な r の選び方は $r=\frac{1}{5}$ であった．従って，r を $\frac{1}{5}$ に近く取るほど (5.7) の収束の速さは遅くなる．

さて，$\{N_2(d)h_{N_2(d)}d^2, d>0\}$ の一様可積分性を示すことにより停止時間 $N_2(d)$ のモーメントに関する次の結果を得る（[42]）．

定理 5.4. f は R 上で有界かつ連続であると仮定する．この時，任意の x に対して
$$N_2(d)h_{N_2(d)}/(b^2f(x)d^{-2})\to1\ \text{a.s.}\ (d\to+0),$$
$$E(N_2(d)h_{N_2(d)})/(b^2f(x)d^{-2})\to1\quad(d\to+0)$$
が成り立つ．さらに，$h_n=n^{-r}$ $(0<r<1)$ とし，q を任意の正数とすると，
$$(N_2(d))^q/(b^2f(x)d^{-2})^{q/(1-r)}\to1\ \text{a.s.}\ (d\to+0),$$
$$E\{(N_2(d))^q\}/(b^2f(x)d^{-2})^{q/(1-r)}\to1\quad(d\to+0)$$
が成り立つ．

ここで，Koronacki and Wertz（[49]）の停止時間を紹介する．K はさらに，$[0,\infty)$ 上で非増加であり，$h_n=cn^{-r}\left(\frac{2}{5}<r<1,\ c>0 \text{ は定数}\right)$ とし，

$$w=\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}h_j^{-1}\int\left\{\int K(z)K\left(z+\frac{y}{h_n}\right)dz\int K(t)K\left(t+\frac{y}{h_j}\right)dt\right\}dy$$

を満たす $0<w<\infty$ が存在するとする．$\{\xi_n, n\geq 1\}$ を $\int f^2(x)dx$ の推定量とする時，任意に与えられた $\alpha\in(0,1)$，正整数 n_0，$\varepsilon>0$ に対して

$$(h_n\xi_n)^{1/2}\leq\frac{nh_n\varepsilon-B}{2(2+r)^{-1/2}w^{1/2}\Phi^{-1}(1-\alpha)}$$

を満たす最小の正整数 $n\geq n_0$ が存在すれば $N(\varepsilon)=n$ とし，そのような n が存在しなければ $N(\varepsilon)=\infty$ として停止時間 $N(\varepsilon)$ を定義する．ここに，B は (5.1) で与えられたものである．

定理 5.5. f は R 上で 2 回連続微分可能で，f, f', f'' は有界であると仮定する．f_n を推定量 WWYE とする時，もし $\xi_n\to\int f^2(x)dx$ a.s. $(n\to\infty)$ ならば

$$\lim_{\varepsilon\to+0}P\left\{\int(f_{N(\varepsilon)}(x)-f(x))^2dx\leq\varepsilon\right\}=1-\alpha$$

が成り立つ．

最後に，タイプⅡの停止時間についてその定義を述べておく．今，$I_n(x)$ を大きさ n の標本を用いて構成された $f(x)$ の信頼区間とし，$|I_n(x)|$ は区間 $I_n(x)$ の長さを表すとする．この時，タイプⅡの停止時間 $N(d)$ を次で定義する．もし $|I_n(x)|\leq 2d$ を満たす最小の正整数 n が存在すれば $N(d)=n$ とし，そのような n が存在しなければ $N(d)=\infty$ とする．

§6. 回帰関数の推定

$Z=(X, Y)$，$\{Z_n=(X_n, Y_n), n\geq 1\}$ は独立で同一分布に従う $R^p\times R$ に値をとる確率ベクトル列で，$X=x$ における Y の回帰関数を $m(x)=E(Y|X=x)$ とし関数型は未知であると仮定する．確率計画模型 (stochastic design model) におけるノンパラメトリック回帰関数の推定問題とは $Z_1, \cdots, Z_n$ に基づいて $m(x)$ を推定することである ([9], [12], [24], [34], [67] などを参照)．今，$W_{nj}(X)=W_{nj}(X; X_1, \cdots, X_n)$ $(1\leq j\leq n)$ を重み付き関数とする時，Stone [77] は $m(x)$ の推定量として

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

を考え，$m_n(x)$ が一致性を持つための $W_n=\{W_{nj}\}$ に関する十分条件を与えた．ところで，1 次元 $(p=1)$ の場合に Nadaraya [56] と Watson [85] はカーネル K と 0 に収束する正数列 $\{h_n, n\geq 1\}$ を用いて次のような非再帰的推定量を考えた．

$$W_{nj}(x)=\frac{K((x-X_j)h_n^{-1})}{\sum_{i=1}^{n}K((x-X_i)h_n^{-1})}.$$

この推定量に関しては，例えば，[22], [31], [33], [60], [62] などを参照されたい．一方，

$$W_{nj}(x)=\frac{h_j^{-p}K((x-X_j)h_j^{-1})}{\sum_{i=1}^{n}h_i^{-p}K((x-X_i)h_i^{-1})}$$

とおくと，$m_n(x)$ は再帰的推定量になる．これらの推定量の統計的性質については，[2], [20], [30], [38], [50], [51] などで議論されている．また，逐次推定に関しても研究されている ([1],

[40], [44], [70] など). この他に, 固定計画模型 (fixed design model) に対しても研究がなされている. これは, $g(x)$ ($x \in R^p$) を未知な関数とする時, $Y_i = g(x_i) + Z_i$ ($i=1, \cdots, n$) ($Z_1, \cdots, Z_n$ は独立で同一分布に従う平均が 0 の確率変数, $x_1, \cdots, x_n$ は与えられた実数) を満たす標本 (x_1, Y_1), $\cdots$, (x_n, Y_n) に基づいて $g(x)$ を推定することである. ([28], [29], [45], [64] など).

§7. おわりに

以上カーネル法を用いたノンパラメトリック確率密度関数の再帰的推定量とその統計的性質および関連した問題について解説した. 推定量 RPE についても非常に多くの興味ある結果が得られているが, ここではほとんど触れなかった. 今後も種々の分野においてカーネル法を用いた推定量の研究が一層進められると思う. 最後に, 漸近展開に基づく高次の漸近理論がノンパラメトリック確率密度関数の推定問題においても展開できるのではないかと思われる.

文献

[1] Aerts, M., and Geertsema, J. C., Bounded length confidence intervals in nonparametric regression, Sequential Anal., **9**(1990), 171-192.

[2] Ahmad, I. A., and Lin, P. E., Nonparametric sequential estimation of a multiple regression function, Bull. Math. Statist., **17**(1976), 63-75.

[3] Anscombe, F. J., Large-sample theory of sequential estimation, Proc. Cambridge Philos. Soc., **48**(1952), 600-607.

[4] Banon, G., Sur un estimateur non paramétrique de la densité de probabilité, Rev. Statist. Appl., **XXIV**(1976), 61-73.

[5] Banon, G., Nonparametric identification for diffusion processes, SIAM J. control and Optimization, **16**(1978), 380-395.

[6] Bickel, P. J., and Rosenblatt, M., On some global measures of the deviations of density function estimates, Ann. Statist., **6**(1973), 1071-1095.

[7] Cacoullos, T., Estimation of a multivariate density, Ann. Inst. Statist. Math., **18**(1966), 179-189.

[8] Carroll, R. J., On sequential density estimation, Z. Wahrsch. Verw. Gebiete, **36**(1976), 137-151.

[9] Cheng, P. E., Applications of kernel regression estimation : A survey, Commun. Statist. Theory and Methods A, **19**(1990), 4103-4134.

[10] Chow, Y. S., and Robbins, H., On the asymptotic theory of fixed-width sequential confidence intervals for the mean, Ann. Math. Statist., **36** (1965), 457-462.

[11] Cline, D. B. H., Optimal kernel estimation of densities, Ann. Inst. Statist. Math., **42**(1990), 287-303.

[12] Collomb, G., Nonparametric regression : An up-to-date Bibliography, Statistics, **16**(1985), 309-324.

[13] Davies, H. I., and Wegman, E. J., Sequential nonparametric density estimation, IEEE Trans. Inform. Theory, IT-**21**(1975), 619-628.

[14] Deheuvels, P., Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre et uniforme presque sûre des estimateurs de la densité, C. R. Acad. Sci. Paris, **278**(1974), 1217-1220.

[15] Deheuvels, P., Estimation sequentielle de la densite, Contrib. Prob. Est. Mat. Ens. Mat. Anal., 1979, 156-169, Univ. of Granada.

[16] Devroye, L., On the pointwise and the integral convergence of recursive kernel estimates of probability densities, Utilitas Math., **15**(1979), 113-128.

[17] Devroye, L., A Course in Density Estimation, Birkhäuser, Boston, 1987.

[18] Devroye, L., The kernel estimate is relatively stable, Probab. Theory Related Fields, **77**(1988), 521-536.

[19] Devroye, L., Asymptotic performance bounds for the kernel estimate, Ann. Statist., **16**(1988), 1162-1179.

[20] Devroye, L., and Wagner, T. J., On the L_1 convergence of kernel estimators of regression functions with applications in discrimination, Z. Wahrsch. Verw. Gebiete, **51**(1980), 15-25.

[21] Devroye, L., and Györfi, L., Nonparametric Density Estimation : The L_1 View, John Wiley, 1985.

[22] Devroye, L., and Krzyżak, A., An equivalence theorem for L_1 convergence of the kernel regression estimate, J. Statist. Plann. Inference, **23** (1989), 71-82.

[23] Epanechnikov, V. A., Nonparametric estimation of a multivariate probability density, Theor. Prob. Appl., **14**(1969), 153-158.

[24] Eubank, R. L., Spline Smoothing and Nonparametric Regression, Marcel Dekker, Inc., New York and Basel, 1988.

[25] Farrell, R. H., Bounded length confidence in-

tervals for the p-point of a distribution function, III, Ann. Math. Statist., **37**(1966), 586-592.

[26] Fukunaga, K., and Hostetler, L. D., The estimation of the gradient of a density function, with applications in pattern recognition, IEEE Trans. Inform. Theory, IT-**21**(1975), 32-40.

[27] Fryer, M. J., A review of some non-parametric methods of density estimation, J. Inst. Math. Appl., **20**(1977), 335-354.

[28] Georgiev, A. A., Consistent nonparametric multiple regression : The fixed design case, J. Multivariate Anal., **25**(1988), 100-110.

[29] Georgiev, A. A., and Greblicki, W., Nonparametric function recovering from noisy observations, J. Statist. Plann. Inference, **13**(1986), 1-14.

[30] Greblicki, W., and Pawlak, M., Necessary and sufficient consistency conditions for a recursive kernel regression estimate, J. Multivariate Anal., **23**(1987), 67-76.

[31] Györfi, L., Recent results on nonparmetric regression estimate and multiple classification, Problems Control Inform. Theory, **10**(1981), 43-52.

[32] Györfi, L., Strong consistent density estimate from ergodic sample, J. Multivariate Anal., **11** (1981), 81-84.

[33] Hall, P., On iterated logarithm laws for linear arrays and nonparametric regression estimators, Ann. Probab., **19**(1991), 740-757.

[34] Härdle, W., Applied Nonparametric Regression, Cambridge Univ. Press, 1990.

[35] Isogai, E., Strong consistency and optimality of a sequential density estimator, Bull. Math. Statist., **19**(1980), 55-69.

[36] Isogai, E., Stopping rules for sequential density estimation, Bull. Math. Statist., **19**(1981), 53-67.

[37] Isogai, E., Strong uniform consistency of recursive kernel density estimators, Sci. Rep. Niigata Univ. Ser. A, **18**(1982), 15-27.

[38] Isogai, E., A class of nonparametric recursive estimators of a multiple regression function, Bull. Inform. Cybernetics, **20**(1983), 33-44.

[39] Isogai, E., Joint asymptotic normality of nonparametric recursive density estimators at a finite number of distinct points, J. Japan Statist. Soc., **14**(1984), 125-135.

[40] Isogai, E., Asymptotic consistency of fixed-width sequential confidence intervals for a multiple regression function, Ann. Inst. Statist. Math., **38**(1986), 69-83.

[41] Isogai, E., The convergence rate of fixed-width sequential confidence intervals for a probability density function, Sequential Anal., **6** (1987), 55-69.

[42] Isogai, E., A note on sequential density estimation, Sequential Anal., **7**(1988), 11-21.

[43] Isogai, E., Nonparametric recursive estimation in stationary Markov processes, Commun. Statist. Theory and Methods A, 18(1989), 1309-1323.

[44] Isogai, E., Nonparametric probability density estimation using recursive kernel estimators, Doctoral Thesis, Univ. of Tsukuba, 1989.

[45] Isogai, E., Nonparametric estimation of a regression function by delta sequences, Ann. Inst. Statist. Math., **42**(1990), 699-708.

[46] Isogai, E., A note on the asymptotic normality of sequential density estimators, Yokohama Math. J., **39**(1992), 115-124.

[47] Kogure, A., Asymptotically optimal cells for a histogram, Ann. Statist., **15**(1987), 1023-1030.

[48] 小暮厚之, ヒストグラムのための最適な級区間：MISE 基準, 数学, **41**(1989), 237-245.

[49] Koronacki, J., and Wertz, W., A global stopping rule for recursive density estimators, J. Statist. Plann. Inference, **20**(1988), 23-39.

[50] Krzyżak, A., and Pawlak, M., Universal consistency results for Wolverton-Wagner regression function estimate with application in discrimination, Problems Control Inform. Theory, **12** (1983), 33-42.

[51] Krzyżak, A., and Pawlak, M., Almost everywhere convergence of a recursive kernel regression function estimate and classification, IEEE Trans. Inform. Theory, IT-**30**(1984), 91-93.

[52] Loève, M., Probability Theory I. 4 th Edition, Springer-Verlag, 1977.

[53] Masry, E., Almost sure convergence of recursive density estimators for stationary mixing processes, Statist. Probab. Lett., **5**(1987), 249-254.

[54] Masry, E., Recursive probability density estimation for weakly dependent stationary processes, IEEE Trans. Inform. Theory, IT-**32**(1986), 254-267.

[55] Masry, E., and Györfi, L., Strong consistency and rates for recursive probability density estimators of stationary processes, J. Multivariate Anal., **22**(1987), 79-93.

[56] Nadaraya, E. A., On estimating regression, Theor. Prob. Appl., **9**(1964), 141-142.

[57] Nadaraya, E. A., Nonparametric Estimation of Probability Densities and Regression Curves, Kluwer Academic Publishers, 1989.

[58] Nguyen, H. T., Density estimation in a continuous-time stationary Markov process, Ann. Statist., **7**(1979), 341-348.

[59] Nguyen, H. T., Recursive nonparametric estimation in stationary Markov processes, Pub. Inst. Statist., Univ. Paris, **29**(1984), 65-84.

[60] Noda, K., Estimation of a regression fucntion by the Parzen kernel-type density estimators,

Ann. Inst. Statist. Math., **28**(1976), 221-234.
[61] Parzen, E., On estimation of a probability density function and mode, Ann. Math. Statist., **33**(1962), 1065-1076.
[62] Pawlak, M., On the almost everywhere properties of the kernel regression estimate, Ann. Inst. Statist. Math., **43**(1991), 311-326.
[63] Prakasa Rao, B. L. S., Nonparametric Functional Estimation, Academic Press, 1983.
[64] Priestley, M. B., and Chao, M. T., Nonparametric function fitting, J. Roy. Statist. Soc. Ser. B, **34**(1972), 385-392.
[65] Rosenblatt, M., Remarks on some nonparametric estimates of a density function, Ann. Math. Statist., **27**(1956), 832-837.
[66] Rosenblatt, M., Curve estimates, Ann. Math. Statist., **42**(1971), 1815-1842.
[67] Roussas, G. G., Nonparametric Functional Estimation and Related Topics, Kluwer Academic Publishers, 1991.
[68] Roussas, G. G., Recursive estimation of the transition distribution fucntion of a Markov process : Asymptotic normality, Statist. Probab. Lett., **11**(1991), 435-447.
[69] Rychlik, Z., The order of approximation in the random central limit theorem, Lect. Notes in Math., No. **656**(1978), Springer, 225-236.
[70] Samanta, M., On sequential estimation of the regression function, Bull. Inform. Cybernetics, **21**(1984), 19-27.
[71] Schuster, E. F., Note on the uniform convergence of density estimates, Ann. Math. Statist., **41**(1970), 1347-1348.
[72] Seheult, A. H., and Quesenberry, C. P., On unbiased estimation of density functions, Ann. Math. Statist., **42**(1971), 1434-1438.
[73] Sen, P. K., and Ghosh, M., On bounded length sequential confidence intervals based on one-sample rank order statistics, Ann. Math. Statist., **42**(1971), 189-203.
[74] Silverman, B. W., Density Estimation for Statistics and Data Analysis, Chapman and Hall, 1986.
[75] Srivastava, R. C., Estimation of probability density function based on random number of observations with applications, Internat. Statist. Rev., **41**(1973), 77-86.
[76] Stein, E. M., Singular integrals and differentiability properties of functions, Princeton Univ. Press, Princeton, New Jersey, 1970.
[77] Stone, C. J., Consistent nonparametric regression, Ann. Statist., 5(1977), 595-645.
[78] Stone, C., An asymptotically optimal window selection rule for kernel density estimates, Ann. Statist., **12**(1984), 1285-1297.
[79] Stute, W., Sequential fixed-width confidence intervals for a nonparametric density function, Z. Wahrsch. Verw. Gebiete, **62**(1983), 113-123.
[80] Takahata, H., Almost sure convergence of density estimators for weakly dependent stationary processes, Bull. Tokyo Gakugei Univ. Ser. IV, **32**(1980), 11-32.
[81] Tapia, R. A., and Thompson, J. R., Nonparametric Probability Density Estimation, Johns Hopkins Univ. Press, Baltimore, Maryland, 1978.
[82] Tran, L. T., Recursive density estimation under dependence, IEEE Trans. Inform. Theory, **35**(1989), 1103-1108.
[83] Tran, L. T., Recursive kernel density estimators under a weak dependence condition, Ann. Inst. Statist. Math., **42**(1990), 305-329.
[84] Watanabe, M., On convergences of asymptotically optimal discriminant functions for pattern classification problems, Bull. Math. Statist., **16** (1974), 23-34.
[85] Watson, G. S., Smooth regression analysis, Sankhyā Ser. A, **26**(1964), 359-372.
[86] Wegman, E. J., Nonparametric probability density estimation : I. A summary of available methods, Technometrics, **14**(1972), 533-546.
[87] Wegman, E. J., and Davies, H. I., Remarks on some recursive estimators of a probability density, Ann. Statist., **7**(1979), 316-327.
[88] Wertz, W., Sequential and recursive estimators of the probability density, Statistics, **16** (1985), 277-295.
[89] Wertz, W., and Schneider, B., Statistical density estimation : a Bibliography, Internat. Statist. Rev., **47**(1979), 155-175.
[90] Wolverton, C. T., and Wagner, T. J., Asymptotically optimal discriminant functions for pattern classification, IEEE Trans. Inform. Theory, IT-**15**(1969), 258-265.
[91] Yamato, H., Sequential estimation of a continuous probability density function and mode, Bull. Math. Statist., **14**(1971), 1-12.
[92] Yoshihara, K., Weakly Dependent Stochastic Sequences and Their Applications Vol. I, Sanseido, 1992.

(1992年6月8日提出)
(いそがい　えいいち・新潟大学理学部)