数理解析研究所講究録
第 879 巻 1994 年 74-86

可算個の値の母数で漸近的に確率１で正しく推定する一致推定量について

筑波大 数学 佐藤道一 (Michikazu Sato)

一致推定量 (弱一致の意味) の収束の速さを考える際に、母数の値を固定して考えると以下の定理２と定理３のように極端なことも起こること (定理１はその言いかえ) に注意する必要がある.

以下、Θ を母数空間, θ を母数、(Y, d) を距離空間, $g: \Theta \to Y$ を写像、δ_n を標本の大きさが n ($i.i.d.$ でなくてもよい) のときの推定量とする.

<u>定理１</u> θ を固定する. このとき、任意の正数列 $\{k_n\}$ に対して $P_\theta\{d(\delta_n, g(\theta)) \leqq 1/k_n\} \not\to 0$ $(n \to \infty)$ となることは、$P_\theta\{\delta_n = g(\theta)\} \not\to 0$ $(n \to \infty)$ と同値である.

<u>証明</u> $P_\theta\{\delta_n = g(\theta)\} \not\to 0$ のとき、

$$P_\theta\{d(\delta_n, g(\theta)) \leqq 1/k_n\} \geqq P_\theta\{\delta_n = g(\theta)\}$$

だから、辺々の $\limsup$ をとり、$P_\theta\{d(\delta_n, g(\theta)) \leqq 1/k_n\} \not\to 0$ が得られる. 逆を示そう. n を固定する.

$$A_k := \{d(\delta_n, g(\theta)) \leqq 1/k\}$$

とおくと、

$$A_1 \supset A_2 \supset \cdots, \quad \bigcap_{k=1}^{\infty} A_k = \{\delta_n = g(\theta)\}$$

だから、

$$\lim_{n\to\infty} P_\theta(A_k) = P_\theta\{\delta_n = g(\theta)\}$$

となり、各nに対してk_nを十分大きくとって

$$P_\theta\{d(\delta_n, g(\theta) \leqq 1/k_n\} \leqq P_\theta\{\delta_n = g(\theta)\} + 1/n$$

だから、辺々の$\limsup$をとり、$P_\theta\{\delta_n = g(\theta)\} \not\to 0$ が得られる。 □

<u>定理2</u>　$g(\theta)$ の一致推定量δ_nが存在すると仮定する。このとき、点列$\{\theta_k\}_{k=1}^{\infty} \subset \Theta$ が与えられれば、

$$\lim_{n\to\infty} P_{\theta_k}\{\delta'_n = g(\theta_k)\} = 1, \quad k = 1, 2, \cdots$$

となる$g(\theta)$ の一致推定量δ'_nが存在する。

<u>証明</u>　各$j, k \in \mathbb{N}$ に対して

$$\lim_{n\to\infty} P_{\theta_k}\{d(\delta_n, g(\theta_k)) \leqq 1/j\} = 1$$

だから、$\nu_{j,k} \in \mathbb{N}$ があって

$$P_{\theta_k}\{d(\delta_n, g(\theta_k)) \leqq 1/j\} \geqq 1 - 1/j \quad \forall n \geqq \nu_{j,k}$$

とできる。$1 = \nu_{1,k} < \nu_{2,k} < \cdots$ としてよい。そこで、$J_{n,k}$ を $\nu_{J_{n,k},k} \leqq n < \nu_{J_{n,k}+1,k}$ で定め、δ'_n を

$$\delta_n' := \begin{cases} g(\theta_1) & d(\delta_n, g(\theta_1)) \leqq 1/J_{n,1} \leqq 1/1 \text{ のとき} \\ g(\theta_2) & \text{上以外で} \\ & d(\delta_n, g(\theta_2)) \leqq 1/J_{n,2} \leqq 1/2 \text{ のとき} \\ g(\theta_3) & \text{上の2つ以外で} \\ & d(\delta_n, g(\theta_3)) \leqq 1/J_{n,3} \leqq 1/3 \text{ のとき} \\ \vdots & \\ \delta_n & \text{その他} \end{cases}$$

で定める。

kを固定すると、十分大きいnに対して

$$1/J_{n,k} \leqq 1/k \qquad (1)$$

であり、$g(\theta_l) \neq g(\theta_k)$, $l < k$ となるlを $l_1, l_2, \cdots, l_m$ とすると、

「$d(\delta_n, g(\theta_k)) \leqq 1/J_{n,k}$

かつ $d(\delta_n, g(\theta_{l_r})) \leqq 1/J_{n,l_r}$

ならば $d(g(\theta_k), g(\theta_{l_r})) \leqq 1/J_{n,k} + 1/J_{n,l_r}$」 (2)

だから、十分大きいnに対しては、(1)(2) より

「$d(\delta_n, g(\theta_k)) \leqq 1/J_{n,k}$ ならば

$d(\delta_n, g(\theta_{l_r})) > 1/J_{n,l_r} \quad \forall r = 1, \cdots, m$

従って $\delta_n' = g(\theta_k)$」 (3)

よって、

$$P_{\theta_k}\{\delta_n' = g(\theta_k)\}$$
$$\geqq P_{\theta_k}\{d(\delta_n, g(\theta_k)) \leqq 1/J_{n,k}\} \quad (n: 十分大)$$
$$\geqq 1 - 1/J_{n,k} \to 1 \ (n \to \infty)$$

である．ここで1番目の不等号は(3)，2番目は $J_{n,k}$ の定義による．

従って、

$$\lim_{n\to\infty} P_{\theta_k}\{\delta_n' = g(\theta_k)\} = 1, \quad k = 1, 2, \cdots$$

は示された．あとは δ_n' が $g(\theta)$ の一致推定量であることを言えばよい．

$\theta \in \Theta$ と $\varepsilon > 0$ を固定する．M を $1/M < \varepsilon/2$ となるようにとり，$N := \max\{M, \nu_{M,1}, \nu_{M,2}, \cdots, \nu_{M,M}\}$ とする．

「$n \geqq N$, $d(\delta_n, g(\theta)) \leqq \varepsilon/2$ ならば $d(\delta_n', g(\theta)) \leqq \varepsilon$」 (4)

を示そう．$\delta_n = \delta_n'$ のときは自明だから、否のときを言う．δ_n' の定義より，

$$d(\delta_n, g(\theta_k)) \leqq 1/J_{n,k} \leqq 1/k \qquad (5)$$

となる最小の k があって，$\delta_n' = g(\theta_k)$ であり、$k < M$ のときは $n \geqq N \geqq \nu_{M,k}$ だから $J_{n,k}$ の定義より $M \leqq J_{n,k}$，よって $1/J_{n,k} \leqq 1/M$ であり，これと (5) から，

$$d(\delta_n, g(\theta_k)) \leqq 1/M \qquad (6)$$

である。$k \geqq M$ ならば (5) から明らかに (6) が成り立つ。従って (6) は任意の k に対して成り立ち、これと M のとり方から、

$$d(\delta_n, g(\theta_k)) \leqq \varepsilon/2 \qquad (7)$$

であり、従って、

$$\begin{aligned} & d(\delta_n', g(\theta)) \\ &= d(g(\theta_k), g(\theta)) && [(5)\text{の直後より}] \\ &\leqq d(g(\theta_k), \delta_n) + d(\delta_n, g(\theta)) \\ &\leqq \varepsilon/2 + \varepsilon/2 = \varepsilon && [\text{仮定と}(7)\text{より}] \end{aligned}$$

である。従って (4) が示された。これより、$n \geqq N$ に対して

$$\begin{aligned} & P_\theta\{d(\delta_n', g(\theta)) \leqq \varepsilon\} \\ & \geqq P_\theta\{d(\delta_n, g(\theta)) \leqq \varepsilon/2\} \to 1 \quad (n \to \infty) \end{aligned}$$

だから、δ_n' は $g(\theta)$ の一致推定量である。 □

<u>定理3</u>　$g(\theta)$ の一致推定量 δ_n が存在すると仮定する。このとき、点列 $\{\eta_k\}_{k=1}^{\infty} \subset Y$ が与えられ、次の正則条件
「各 k に対して $\{A_{k,\ell}\}_{\ell=1}^{\infty}$ があって、
$\{\theta : g(\theta) = \eta_k\} = \bigcup_{\ell=1}^{\infty} A_{k,\ell}$ で δ_n は各 $A_{k,\ell}$ 上での $g(\theta)$ の一様一致推定量である。」
が満たされると仮定する。このとき、$g(\theta) = \eta_k$ となる k が存在するようなすべての θ に対して

$$\lim_{n\to\infty} P_\theta\{\delta_n = g(\theta)\} = 1$$

となる $g(\theta)$ の一致推定量 δ_n' が存在する。

証明　$A_{k,l} \neq \phi$ としてよい。番号をつけかえて、$\{A_{k,l} : k, l \in \mathbb{N}\} = \{B_k : k \in \mathbb{N}\}$ とする。各 j, k に対して

$$\lim_{n\to\infty} P_\theta\{d(\delta_n, g(\theta)) \leqq 1/j\} = 1$$
$$(\theta \in B_k \text{ 上で一様に})$$

だから、$\nu_{j,k} \in \mathbb{N}$ (θ によらない) があって、

$$P_\theta\{d(\delta_n, g(\theta)) \leqq 1/j\} \geqq 1 - 1/j \quad \forall n \geqq \nu_{j,k}\ \forall \theta \in B_k$$

であり、あとは $\theta_k \in B_k$ として定理2と同様に証明できる。

□

注意　(i) 定理3の正則条件は、特に「$\Theta = \bigcup_{l=1}^{\infty} K_l$ で、δ_n は各 K_l 上での $g(\theta)$ の一様一致推定量である。」が成り立てば満たされる。これは特に「Θ は σコンパクトで δ_n は $g(\theta)$ の局所一様一致推定量である。」が成り立てば満たされる。

(ii) 正則条件を省いてはならない (後述の例1、例2)。

(iii) 別な正則条件の下で、$\{\eta_k\}_{k=1}^{\infty}$ を非可算無限個の値に変えてはならない (次の定理4)。

定理4　(i) (小標本論) X の分布は θ によらないある σ有限測度 ν に関して絶対連続とすると、推定量 $\delta = \delta(X)$ につ

いて、$\{g(\theta) : P_\theta\{\delta = g(\theta)\} \neq 0\}$ は高々可算である。

(ii)（大標本論）各nに対して、$(X_1, \cdots, X_n)$の分布はθによらないあるσ有限測度ν_nに関して絶対連続とすると、推定量$\delta_n = \delta_n(X_1, \cdots, X_n)$ について、

$\{g(\theta) : P_\theta\{\delta_n = g(\theta)\} \nrightarrow 0 \ (n \to \infty)\}$ は高々可算である。

<u>証明</u> (i) νをとり直して確率測度としてよい。$S := \{g(\theta) : P_\theta\{\delta = g(\theta)\} \neq 0\}$ とする。$\eta \in S$とすると、θがあって$g(\theta) = \eta$, よって $P_\theta\{\delta = \eta\} \neq 0$, よって$\nu\{\delta = \eta\} \neq 0$である。$\nu$は確率測度だからそのような$\eta$は高々可算個である。(ii) は (i) から明らか。 □

<u>例1</u>（i.i.d.でない例）Θ を $\{2^{-1}, 2^{-2}, \cdots\}$ の値をとる収束数列全体とし、$\theta = \{\theta_n\}_{n=1}^{\infty} \in \Theta$ に対して $P_\theta\{X_n = \theta_n\} = 1$ とし、$g(\theta) := \lim_{n\to\infty} \theta_n$ の推定を考える。$\delta_n(X_1, \cdots, X_n) = X_n$ は $g(\theta)$ の強一致推定量であるが、$(X_1, \cdots, X_n)$ に基づく確率化推定量R_n（すなわち、R_nは確率変数で、$(X_1, \cdots, X_n)$ の値を与えたときのR_nの条件付き分布はθに依存しない）で、R_nは$g(\theta)$の一致推定量で、

$$g(\theta) = 0 \Longrightarrow \liminf_{n\to\infty} P_\theta\{R_n = g(\theta)\} > 0 \quad (8)$$

となるものは存在しない。

[存在したとする．$j=1, 2, \cdots$ に対して，
$2^{-j}-2^{-j-2}<R_n<2^{-j}+2^{-j+2}$ のときは $R_n=2^{-j}$
と定義し直すと，これも条件を満たし，

$$g(\theta)=2^{-j} \Rightarrow \lim_{n\to\infty} P_\theta\{R_n=2^{-j}\}=1$$

である．

すると，まず $\theta^{(1)}=\{2^{-1}, 2^{-1}, \cdots\}$ に対して ν_1 があって

$$P_{\theta^{(1)}}\{R_n=2^{-1}\}>1-2^{-1} \quad \forall n \geqq \nu_1$$

であり，特に

$$P_{\theta^{(1)}}\{R_{\nu_1}=2^{-1}\}>1-2^{-1}$$

であり，今 $P_{\theta^{(1)}}\{X_1=\cdots=X_{\nu_1}=2^{-1}\}=1$ だから，

$$P_{\theta^{(1)}}\{R_{\nu_1}=2^{-1} \mid X_1=\cdots=X_{\nu_1}=2^{-1}\}>1-2^{-1}$$

である．今，$P_\theta\{A|B\}$ が θ によらない場合 $P\{A|B\}$ と書くことにすると，R_{ν_1} は $X_1, \cdots, X_{\nu_1}$ に基づく確率化推定量だから，

$$P\{R_{\nu_1}=2^{-1} \mid X_1=\cdots=X_{\nu_1}=2^{-1}\}>1-2^{-1}$$

である．次に $\theta^{(2)}=\{2^{-1}, \cdots, \underset{\substack{\uparrow\\ \text{第}\nu_1\text{項}}}{2^{-1}}, 2^{-2}, 2^{-2}, \cdots\}$ に対して ν_2 があって，

$$P_{\theta^{(2)}}\{R_n=2^{-2}\}>1-2^{-2} \quad \forall n \geqq \nu_2$$

である．$\nu_1<\nu_2$ としてよい．特に

$$P_{\theta^{(2)}}\{R_{\nu_2}=2^{-2}\}>1-2^{-2}$$

であり、

$$P_{\theta^{(2)}}\left\{\begin{matrix} X_1=\cdots=X_{\nu_1}=2^{-1} \\ X_{\nu_1+1}=\cdots=X_{\nu_2}=2^{-2}\end{matrix}\right\}=1$$

だから、

$$P_{\theta^{(2)}}\left\{R_{\nu_2}=2^{-2}\,\middle|\,\begin{matrix} X_1=\cdots=X_{\nu_1}=2^{-1} \\ X_{\nu_1+1}=\cdots=X_{\nu_2}=2^{-2}\end{matrix}\right\}>1-2^{-2}$$

で、前と同様に、

$$P\left\{R_{\nu_2}=2^{-2}\,\middle|\,\begin{matrix} X_1=\cdots=X_{\nu_1}=2^{-1} \\ X_{\nu_1+1}=\cdots=X_{\nu_2}=2^{-2}\end{matrix}\right\}>1-2^{-2}$$

である。同じことをくり返して、$\nu_1<\nu_2<\cdots$ があって、

$$P\left\{R_{\nu_j}=2^{-j}\,\middle|\,\begin{matrix} X_1=\cdots=X_{\nu_1}=2^{-1} \\ X_{\nu_1+1}=\cdots=X_{\nu_2}=2^{-2} \\ \cdots\cdots\cdots\cdots \\ X_{\nu_{j-1}+1}=\cdots=X_{\nu_j}=2^{-j}\end{matrix}\right\}>1-2^{-j}$$

$$j=1,2,3,\cdots$$

となる。従って、

$$\theta^*=\{2^{-1},\cdots,\underset{\substack{\uparrow\\ \text{第}\nu_1\text{項}}}{2^{-1}},2^{-2},\cdots,\underset{\substack{\uparrow\\ \text{第}\nu_2\text{項}}}{2^{-2}},2^{-3},\cdots,\underset{\substack{\uparrow\\ \text{第}\nu_3\text{項}}}{2^{-3}},\cdots\}$$

に対しては、$g(\theta^*)=0$ で、

$$P_{\theta^*}\{R_{\nu_j}=2^{-j}\}>1-2^{-j}$$

従って、

$$P_{\theta^*}\{R_{\nu_j}=0\} \leqq 2^{-j}$$

よって、

$$\liminf_{n\to\infty} P_{\theta^*}\{R_n=0\}=0$$

となり矛盾である。]

なお、この例では (8) で liminf を limsup にしてはならない。もっと詳しくは、非確率化推定量 δ_n' で、$g(\theta)$の強一致推定量であって、

$g(\theta)=0 \Rightarrow$ 無限個の n に対して $P_\theta\{\delta_n'=g(\theta)\}=1$

となるものが存在する。

[$$\delta_n' := \begin{cases} 0 & X_1>X_n,\ X_2>X_n,\ \cdots,\ X_{n-1}>X_n \text{のとき} \\ \delta_n\ (=X_n) & \text{その他} \end{cases}$$

とする。$\delta_n=\theta_n$ としてよい。$g(\theta)\neq 0$ なら十分大きい n に対して $\delta_n=g(\theta)$, 従って $\delta_n'=\delta_n\to g(\theta)$ であり, $g(\theta)=0$ なら $\theta_n>0$, $\theta_n\to 0$ だから、$\{\delta_n'\}$ は(定数列で)無限個の n に対して 0 となる。すなわち無限個の n に対して $P_\theta\{\delta_n=g(\theta)\}=1$ で、また $|\delta_n'|\leqq\delta_n=\theta_n\to 0$ よって $\delta_n'\to 0$ である。]

例2 (i.i.d. の例) Θ, $g(\theta)$ を例1と同じものとし、$K_1, K_2, \cdots$ を i.i.d. で $P_\theta\{K_n=k\}=2^{-k}$ ($k=1, 2, 3, \cdots$) とし、$P_\theta\{Y_n=\theta_k \mid K_n=k\}=1$, $X_n=(K_n, Y_n)$ とする。このとき $X_1, X_2, \cdots$ は i.i.d.

であり、$\max\{K_1, \cdots, K_n\} = K_{J_n} = K_n^*$ とすると、

$\delta_n(X_1, \cdots, X_n) = Y_{J_n}$ は $g(\theta)$ の強一致推定量である。

[各 k に対して $P\{K_n \leqq k \ \exists n\} = 1$ だから、P_θ-a.e. で $K_n^* \to \infty$、また $K_1^* \leqq K_2^* \leqq \cdots$ であり、よって $\delta_n = Y_{J_n} = \theta_{K_n^*} \to g(\theta)$ である。]

しかし例1の条件 (8) を満たす $g(\theta)$ の（確率化）一致推定量 R_n は存在しない。

[存在したとする。V を $(X_1, X_2, \cdots)$ と独立で $(0, 1)$ 上の一様分布に従う確率変数とし（これは実際に確率化する手続きである。）R_n は $(X_1, \cdots, X_n, V)$ の函数としてよい。従って R_n は $(K_1, \cdots, K_n, \theta_1, \cdots, \theta_{K_n}, V)$ の函数としてよい。そこで、

$$R_k^* := \begin{cases} R_k & K_n \leqq k \quad \forall n = 1, \cdots, k \text{ のとき} \\ 0 & \text{その他} \end{cases}$$

とすると（「その他」は実は本質的ではない。）、R_k^* は $(K_1, \cdots, K_k, V, \theta_1, \cdots, \theta_k)$ の函数で、$(K_1, \cdots, K_k, V)$ の分布は θ によらないので R_k^* は $(\theta_1, \cdots, \theta_k)$ に基づく確率化推定量（従って例1での推定量）で、

$$P\{K_n \leqq k \quad \forall n = 1, \cdots, k\}$$
$$= (P\{K_1 \leqq k\})^k$$

$$=\left\{\frac{\frac{1}{2}\left(1-\frac{1}{2^{k}}\right)}{1-\frac{1}{2}}\right\}^{k}=\left(1-\frac{1}{2^{k}}\right)^{k}\to 1\ (k\to\infty) \qquad (9)$$

であり、一般のθに対して、

$$P_\theta\{|R_k^*-g(\theta)|>\varepsilon\}$$
$$=P_\theta\{|R_k^*-g(\theta)|>\varepsilon \text{ かつ } K_n\leqq k\ \ \forall n=1,\cdots,k\}$$
$$+P_\theta\{|R_k^*-g(\theta)|>\varepsilon \text{ かつ非「} K_n\leqq k\ \ \forall n=1,\cdots,k\text{」}\}$$
$$\leqq P_\theta\{|R_k-g(\theta)|>\varepsilon\}+P\{\text{非「}K_n\leqq k\ \ \forall n=1,\cdots,k\text{」}\}$$
$$=2-P_\theta\{|R_k-g(\theta)|\leqq\varepsilon\}-P\{K_n\leqq k\ \ \forall n=1,\cdots,k\}$$
$$\to 0\ (k\to\infty)\ \ (R_k\text{の一致性と(9)による.})$$

だから、R_k^*は$g(\theta)$の一致推定量で、また、$g(\theta)=0$となるθに対しては、上と同様の計算により、

$$P_\theta\{R_k^*\neq g(\theta)\}$$
$$\leqq 2-P_\theta\{R_k=g(\theta)\}-P\{K_n\leqq k\ \ \forall n=1,\cdots,k\}$$

だから、

$$P_\theta\{R_k^*=g(\theta)\}$$
$$\geqq P_\theta\{R_k=g(\theta)\}+P\{K_n\leqq k\ \ \forall n=1,\cdots,k\}-1$$

であり、辺々の$\liminf$をとると、R_kが(8)を満たすことと(9)により、

$$\liminf_{k\to\infty} P_\theta\{R_k^*=g(\theta)\}\geqq\liminf_{k\to\infty} P_\theta\{R_k=g(\theta)\}>0$$

となり例1に矛盾する。]

なお$\liminf$を$\limsup$にしてよいかどうかは不明である。

注意　Lehmann (1983, p. 335) には，δ_n を $g(\theta)$ の一致推定量とすると，$\{k_n\}$ が十分速く ∞ に発散すれば $P_\theta\{|\delta_n - g(\theta)| \leqq a/k_n\} \to 0$ となると書かれているが，定理1，定理2よりこれは誤りである．

参考文献

Lehmann, E. L. (1983). *Theory of Point Estimation.* John Wiley & Sons, Inc., New York.