

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 はその言い換え) 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、 $g: \Theta \rightarrow Y$ 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。

定理 1 θ を固定する。このとき、任意の正数列 $\{k_n\}$ に対して $P_\theta \{d(\delta_n, g(\theta)) \leq 1/k_n\} \rightarrow 0 \ (n \rightarrow \infty)$ となることは、 $P_\theta \{\delta_n = g(\theta)\} \rightarrow 0 \ (n \rightarrow \infty)$ と同値である。

証明 $P_\theta \{\delta_n = g(\theta)\} \rightarrow 0$ のとき、

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

だから、辺々の $\limsup$ をとり、 $P_\theta \{d(\delta_n, g(\theta)) \leq 1/k_n\} \rightarrow 0$ が得られる。逆を示そう。 n を固定する。

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

とかくと、

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

だから、

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

となり、各 n に対して k_n を十分大きくとって

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

だから、辺々の $\limsup$ をとり、 $P_{\theta}\{\delta_n = g(\theta)\} \rightarrow 0$ が得られる。 $\square$

定理 2 $g(\theta)$ 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、

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

となる $g(\theta)$ の一致推定量 δ'_n が存在する。

証明 各 $j, k \in \mathbb{N}$ に対して

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

だから、 $\nu_{j,k} \in \mathbb{N}$ があって

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

とできる。 $1 = \nu_{1,k} < \nu_{2,k} < \dots$ としてよい。そこで、 $J_{n,k}$

を $\nu_{J_{n,k}, k} \leq n < \nu_{J_{n,k}+1, k}$ で定め、 δ'_n を

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

で定める。

k を固定すると、十分大きい n に対して

$$1/J_{n,k} \leq 1/k \quad (1)$$

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

$$\begin{aligned} & \text{「} d(\delta_n, g(\theta_k)) \leq 1/J_{n,k} \\ & \text{かつ } d(\delta_n, g(\theta_{l_r})) \leq 1/J_{n,l_r} \\ & \text{ならば } d(g(\theta_k), g(\theta_{l_r})) \leq 1/J_{n,k} + 1/J_{n,l_r} \text{」} \end{aligned} \quad (2)$$

だから、十分大きい n に対しては、(1)(2) より

$$\begin{aligned} & \text{「} d(\delta_n, g(\theta_k)) \leq 1/J_{n,k} \text{ ならば} \\ & d(\delta_n, g(\theta_{l_r})) > 1/J_{n,l_r} \quad \forall r = 1, \dots, m \end{aligned}$$

$$\text{従って } \delta'_n = g(\theta_k) \quad (3)$$

よって、

$$\begin{aligned} & P_{\theta_k} \{ \delta'_n = g(\theta_k) \} \\ & \geq P_{\theta_k} \{ d(\delta_n, g(\theta_k)) \leq 1/J_{n,k} \} \quad (n: \text{十分大}) \\ & \geq 1 - 1/J_{n,k} \rightarrow 1 \quad (n \rightarrow \infty) \end{aligned}$$

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、

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

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 $1/M < \varepsilon/2$ とする
ようにとり、 $N := \max \{ M, \nu_{M,1}, \nu_{M,2}, \dots, \nu_{M,M} \}$ とする。

$$\text{「 } n \geq N, d(\delta_n, g(\theta)) \leq \varepsilon/2 \text{ ならば } d(\delta'_n, g(\theta)) \leq \varepsilon \text{」} \quad (4)$$

を示そう。 $\delta_n = \delta'_n$ のときは自明なから、否のときを言う。
 δ'_n の定義より、

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

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 $n \geq N \geq \nu_{M,k}$ だから $J_{n,k}$ の定義より $M \leq J_{n,k}$ 、
よって $1/J_{n,k} \leq 1/M$ であり、これと(5)から、

$$d(\delta_n, g(\theta_k)) \leq 1/M \quad (6)$$

である。 $k \geq M$ 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、

$$d(\delta_n, g(\theta_k)) \leq \varepsilon/2 \quad (7)$$

であり、従って、

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

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

$$\begin{aligned} P_\theta \{d(\delta'_n, g(\theta)) \leq \varepsilon\} \\ \geq P_\theta \{d(\delta_n, g(\theta)) \leq \varepsilon/2\} \rightarrow 1 \quad (n \rightarrow \infty) \end{aligned}$$

だから、 δ'_n は $g(\theta)$ の一致推定量である。 □

定理3 $g(\theta)$ 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 $\{A_{k,l}\}_{l=1}^\infty$ があって、
 $\{\theta: g(\theta) = \eta_k\} = \bigcup_{l=1}^\infty A_{k,l}$ で δ_n は各 $A_{k,l}$ 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

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

となる $g(\theta)$ の一致推定量 δ_n' が存在する。

証明 $A_{k,l} \neq \emptyset$ 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

$$\lim_{n \rightarrow \infty} P_\theta \{ d(\delta_n, g(\theta)) \leq 1/j \} = 1$$

($\theta \in B_k$ 上で一様に)

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

$P_\theta \{ d(\delta_n, g(\theta)) \leq 1/j \} \geq 1 - 1/j \quad \forall n \geq \nu_{j,k} \quad \forall \theta \in B_k$
であり、あとは $\theta \in B_k$ 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「 $\Theta = \bigcup_{k=1}^{\infty} K_k$ で、 δ_n は各 K_k 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、 $(X_1, \dots, X_n)$ の分布は θ によらないある有限測度 ν_n に関して絶対連続とすると、推定量 $\delta_n = \delta_n(X_1, \dots, X_n)$ について、

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

証明 (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。 ν は確率測度だからそのような η は高々可算個である。(ii) は (i) から明らか。 $\square$

例 1 (i.i.d. でない例) Θ を $\{2^{-1}, 2^{-2}, \dots\}$ 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、 $g(\theta) := \lim_{n \rightarrow \infty} \theta_n$ の推定を考える。 $\delta_n(X_1, \dots, X_n) = X_n$ は $g(\theta)$ の強一致推定量であるが、 $(X_1, \dots, X_n)$ に基づく確率化推定量 R_n (すなわち、 R_n は確率変数で、 $(X_1, \dots, X_n)$ 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、

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

となるものは存在しない。

[存在したとする. $j=1, 2, \dots$ 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、

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

である。

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

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

であり、特に

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

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

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

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

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

である。次に $\theta^{(2)} = \{2^{-1}, \dots, \underset{\substack{\uparrow \\ \text{第 } \nu_1 \text{ 項}}}{2^{-1}}, 2^{-2}, 2^{-2}, \dots\}$ に対し

て ν_2 があって、

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

である。 $\nu_1 < \nu_2$ としてよい。特に

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

$$P_{\theta^*} \{R_{1/2} = 0\} \leq 2^{-2}$$

よって、

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

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、

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

となるものが存在する。

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

とする。 $\delta_n = \theta_n$ としてよい。 $g(\theta) \neq 0$ ならば十分大きい n に対して $\delta_n = g(\theta)$ 、従って $\delta'_n = \delta_n \rightarrow g(\theta)$ であり、 $g(\theta) = 0$ ならば $\theta_n > 0$ 、 $\theta_n \rightarrow 0$ だから、 $\{\delta'_n\}$ は (定数数列で) 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 $|\delta'_n| \leq \delta_n = \theta_n \rightarrow 0$ よって $\delta'_n \rightarrow 0$ である。]

例 2 (i. i. d. の例) ④, $g(\theta)$ を例 1 と同じものとし、 $K_1, K_2, \dots$ を i. i. d. で $P_{\theta} \{K_n = k\} = 2^{-k}$ ($k=1, 2, 3, \dots$) とし、 $P_{\theta} \{Y_n = \theta_k | K_n = k\} = 1$, $X_n = (K_n, Y_n)$ とする。このとき $X_1, X_2, \dots$ は i. i. d.

であり、 $\max\{K_1, \dots, K_n\} = K_{J_n} = K_n^*$ とすると、
 $\delta_n(X_1, \dots, X_n) = Y_{J_n}$ は $g(\theta)$ の強一致推定量である。
 [各 k に対して $P\{K_n \leq k \mid \exists n\} = 1$ だから、 P_θ -a.e. で $K_n^* \rightarrow \infty$ 、また $K_1^* \leq K_2^* \leq \dots$ であり、よって $\delta_n = Y_{J_n} = \theta_{K_n^*} \rightarrow g(\theta)$ である。]

しか(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 $(X_1, X_2, \dots)$ 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 $(X_1, \dots, X_n, V)$ の函数としてよい。従って R_n は $(K_1, \dots, K_n, \theta_1, \dots, \theta_{K_n}, V)$ の函数としてよい。そこで、

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

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

$$\begin{aligned} & P\{K_n \leq k \mid \forall n = 1, \dots, k\} \\ &= (P\{K_1 \leq k\})^k \end{aligned}$$

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

であり、一般の θ に対して、

$$\begin{aligned} & P_\theta \{ |R_k^* - g(\theta)| > \varepsilon \} \\ &= P_\theta \{ |R_k^* - g(\theta)| > \varepsilon \text{ かつ } K_n \leq k \quad \forall n=1, \dots, k \} \\ &\quad + P_\theta \{ |R_k^* - g(\theta)| > \varepsilon \text{ かつ 非 } \lceil K_n \leq k \quad \forall n=1, \dots, k \rceil \} \\ &\leq P_\theta \{ |R_k - g(\theta)| > \varepsilon \} + P \{ \text{非 } \lceil K_n \leq k \quad \forall n=1, \dots, k \rceil \} \\ &= 2 - P_\theta \{ |R_k - g(\theta)| \leq \varepsilon \} - P \{ K_n \leq k \quad \forall n=1, \dots, k \} \\ &\rightarrow 0 \quad (k \rightarrow \infty) \quad (R_k \text{ の一致性と (9) による.}) \end{aligned}$$

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、

$$\begin{aligned} & P_\theta \{ R_k^* \neq g(\theta) \} \\ &\leq 2 - P_\theta \{ R_k = g(\theta) \} - P \{ K_n \leq k \quad \forall n=1, \dots, k \} \end{aligned}$$

だから、

$$\begin{aligned} & P_\theta \{ R_k^* = g(\theta) \} \\ &\geq P_\theta \{ R_k = g(\theta) \} + P \{ K_n \leq k \quad \forall n=1, \dots, k \} - 1 \end{aligned}$$

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、

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

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 $P_\theta\{|\delta_n - g(\theta)| \leq a/k_n\} \rightarrow 0$ 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.