数理解析研究所講究録
1101 巻 1999 年 38-42

予測十分性と逐次推測

筑波大・数学　西平 祐治 (Yuji Nishihira)
筑波大・数学　赤平 昌文 (Masafumi Akahira)

1 はじめに

統計的決定論において, 予測十分性 (prediction sufficiency) の概念は, 通常の十分性 (sufficiency) を統計的予測論に拡張して作られたものである.

まず, Xを観測される確率変数 (データ) とし, Yを未観測の確率変数としたとき, Xに基づいてYを予測する問題を考える. そこで, $\mathcal{P}$を (X,Y) の確率分布の族として, Xに基づく統計量 $T=T(X)$ がYに対して予測十分であるという概念を, Tが Xの周辺確率分布族 $\mathcal{P}^X$ に対して十分であるという条件に, さらに Tを与えたときに Xと Yが条件付独立であるという条件を付加することによってつくることができる. そして, この予測十分性の概念が十分性の場合と同様に Neyman 型の因子分解定理, リスクによる特徴付け等の性質をもつことはすでに知られている ([S67], [SM69], [TA75], [T77], [A82]).

本論では, 測度論的設定の下で部分σ-加法族を用いて予測十分性の概念について考察し, また, 逐次の場合に, Bahadur[B54] における推移性 (transitivity) との関連についても論じる ([AN99]).

2 予測十分性

まず, $(\mathcal{X},\mathcal{A})$ を標本空間, $\mathcal{P}$を $\mathcal{A}$ 上の確率分布族とする. また, $\mathcal{B},\mathcal{C}$を $\mathcal{A}$ の部分σ-加法族, $\mathcal{B}_0$を $\mathcal{B}$の部分σ-加法族とする. 集合 A の定義関数をχ_Aで表わす.

定義 2.1. $\mathcal{B}_0$が $\mathcal{P}$に関して $\mathcal{B}$に対して十分 (sufficient) であるとは, 任意の $B\in\mathcal{B}$に対して, $\mathcal{B}_0$-可測関数$\phi_B^{\mathcal{B}_0}$が存在して, 任意の $p\in\mathcal{P}$に対して

$$\phi_B^{\mathcal{B}_0}=E_p[\chi_B|\mathcal{B}_0]\quad \text{a.e.}[p]$$

が成り立つことである. これを記号で, $\mathcal{B}_0$ suf $(\mathcal{P};\mathcal{B})$ と表わす. ここで $E_p[\chi_B|\mathcal{B}_0]$ は, $\mathcal{B}_0$が与えられたときのχ_Bの p の下での条件付期待値である.

定義 2.2. $\mathcal{B}_0$が $\mathcal{P}$に関して $\mathcal{B}$に対して必要 (necessary) であるとは, 任意の $\mathcal{B}_1$ suf $(\mathcal{P};\mathcal{B})$ に対して

$$\mathcal{B}_0\subset\mathcal{B}_1\ \ (\mathcal{B},\mathcal{P})$$

が成り立つことである．これを記号で, $\mathcal{B}_0$ nec $(\mathcal{P};\mathcal{B})$ と表わす．ここで $\mathcal{B}_0 \subset \mathcal{B}_1\ (\mathcal{B},\mathcal{P})$ は, 任意の $B_0 \in \mathcal{B}_0$に対して, $B_1 \in \mathcal{B}_1$が存在して, 任意の $p \in \mathcal{P}$に対して, $p(B_0 \triangle B_1) = 0$ となることである．ただし, $\triangle$ は2つの集合の対称差を表わす.

次に, $\mathcal{A}$ の部分σ-加法族 $\mathcal{A}_1$, $\mathcal{A}_2$に対して, $\mathcal{A}_1$, $\mathcal{A}_2$を含む最小のσ-加法族を $\mathcal{A}_1 \vee \mathcal{A}_2$によって表わす.

定義 2.3. $(\mathcal{B}_0;\mathcal{B},\mathcal{C})$ が $\mathcal{P}$に対して Markov であるとは, 任意の $C \in \mathcal{C}$と任意の $p \in \mathcal{P}$に対して

$$E_p[\chi_C|\mathcal{B}] = E_p[\chi_C|\mathcal{B}_0] \quad \text{a.e.}[p]$$

が成り立つことである.

このとき, 次の (i), (ii), (iii), (iv) は互いに同値であることが知られている.

(i) $(\mathcal{B}_0;\mathcal{B},\mathcal{C})$ が $\mathcal{P}$に対して Markov である.

(ii) $\mathcal{B}_0$が与えられたとき$\mathcal{B}$と$\mathcal{C}$が条件付独立である, すなわち任意の $B \in \mathcal{B}$, 任意の $C \in \mathcal{C}$, 任意の $p \in \mathcal{P}$に対して

$$E_p[\chi_{B\cap C}|\mathcal{B}_0] = E_p[\chi_B|\mathcal{B}_0]E_p[\chi_C|\mathcal{B}_0] \quad \text{a.e.}[p]$$

である.

(iii) 任意の $B \in \mathcal{B}$, 任意の $C \in \mathcal{C}$, 任意の $p \in \mathcal{P}$に対して

$$E_p[\chi_{B\cap C}|\mathcal{B}_0 \vee \mathcal{C}] = \chi_C E_p[\chi_B|\mathcal{B}_0] \quad \text{a.e.}[p]$$

である.

(iv) 任意の $B \in \mathcal{B}$, 任意の $p \in \mathcal{P}$に対して

$$E_p[\chi_B|\mathcal{C}] = E_p[E_p[\chi_B|\mathcal{B}_0]|\mathcal{C}] \quad \text{a.e.}[p]$$

である.

ここで, 上記の同値性については, “(i)⇔(ii)”は Loève[L63], “(i)⇔(iii)”は Akahira and Takeuchi[AT80], “(i)⇔(iv)”は [B54] による.

定義 2.4. $\mathcal{B}_0$が$\mathcal{C}$と $\mathcal{P}$に関して$\mathcal{B}$に対して予測十分 (prediction sufficient) であるとは, 次の (a), (b) が成り立つことである．これを記号で, $\mathcal{B}_0$ pred.suf $(\mathcal{P};\mathcal{B},\mathcal{C})$ と表わす.

(a) $\mathcal{B}_0$ suf $(\mathcal{P};\mathcal{B})$,

(b) $(\mathcal{B}_0;\mathcal{B},\mathcal{C})$ が $\mathcal{P}$に対して Markov である.

なお, Skibinsky[S67] は予測十分性を adequacy という言葉で呼んでいる.

定理 2.1([AT80]). $\mathcal{S}$を $\mathcal{B}_0 \vee \mathcal{C}$の部分$\sigma$-加法族とする. このとき, $\mathcal{B}_0$ pred.suf $(\mathcal{P};\mathcal{B},\mathcal{C})$ かつ $\mathcal{S}$ suf $(\mathcal{P};\mathcal{B}_0 \vee \mathcal{C})$ ならば, $\mathcal{S}$ suf $(\mathcal{P};\mathcal{B} \vee \mathcal{C})$ である.

上記の定理を用いれば, 次のことが成り立つ.

定理 2.2. $\mathcal{B}_0$ pred.suf $(\mathcal{P};\mathcal{B},\mathcal{C})$ かつ $\mathcal{B}_1$ nec $(\mathcal{P};\mathcal{B} \vee \mathcal{C})$ ならば, $\mathcal{B}_0$ pred.suf $(\mathcal{P};\mathcal{B},\mathcal{B}_1)$ である.

証明. まず仮定より $(\mathcal{B}_0;\mathcal{B},\mathcal{C})$ が $\mathcal{P}$に対して Markov であるので, 任意の $B_0 \in \mathcal{B}_0$, 任意の $C \in \mathcal{C}$, 任意の $p \in \mathcal{P}$に対して

$$\begin{aligned} E_p[\chi_{B_0 \cap C}|\mathcal{B}] &= \chi_{B_0} E_p[\chi_C|\mathcal{B}] \\ &= \chi_{B_0} E_p[\chi_C|\mathcal{B}_0] \\ &= E_p[\chi_{B_0 \cap C}|\mathcal{B}_0] \quad \text{a.e.}[p] \end{aligned}$$

が成り立つ. よって $(\mathcal{B}_0;\mathcal{B},\mathcal{B}_0 \vee \mathcal{C})$ は $\mathcal{P}$に対して Markov である. また, 定理 2.1 より $\mathcal{B}_0 \vee \mathcal{C}$ suf $(\mathcal{P};\mathcal{B} \vee \mathcal{C})$ が成り立つ. ここで, $\mathcal{B}_0 \vee \mathcal{C}$は $\mathcal{P}$に関して $\mathcal{B} \vee \mathcal{C}$に対して十分であり, $\mathcal{B}_1$は仮定より $\mathcal{P}$に関して $\mathcal{B} \vee \mathcal{C}$に対して必要であるので,

$$\mathcal{B}_1 \subset \mathcal{B}_0 \vee \mathcal{C} \ (\mathcal{B} \vee \mathcal{C}, \mathcal{P})$$

が成り立つ. よって, 任意の $B_1 \in \mathcal{B}_1$に対して, $A_0 \in \mathcal{B}_0 \vee \mathcal{C}$が存在して, 任意の $p \in \mathcal{P}$に対して

$$p(B_1 \triangle A_0) = 0$$

が成り立つ. すなわち, 任意の $p \in \mathcal{P}$に対して

$$\chi_{B_1} = \chi_{A_0} \quad \text{a.e.}[p]$$

が成り立つ. このとき

$$\begin{aligned} E_p[\chi_{B_1}|\mathcal{B}] &= E_p[\chi_{A_0}|\mathcal{B}] \\ &= E_p[\chi_{A_0}|\mathcal{B}_0] \\ &= E_p[\chi_{B_1}|\mathcal{B}_0] \quad \text{a.e.}[p] \end{aligned}$$

となるので, $(\mathcal{B}_0;\mathcal{B},\mathcal{B}_1)$ は $\mathcal{P}$に対して Markov である. また $\mathcal{B}_0$ suf $(\mathcal{P};\mathcal{B})$ は仮定より明らかであるので, $\mathcal{B}_0$ pred.suf $(\mathcal{P};\mathcal{B},\mathcal{B}_1)$ が成り立つ. □

3 逐次推定における予測十分性

Bahadur[B54] は, 推移性という概念を導入し, 十分性でかつ推移性の特徴付けを論じた. 最近, Nishihira[N99] も, それに関連した結果を得ている. ここでは, それらと予測十分性との関係について述べる.

まず, $\{\mathcal{A}^{(n)}\}$ を $\mathcal{A}$ の部分σ-加法族の列で, $\mathcal{A}^{(1)} \subset \mathcal{A}^{(2)} \subset \cdots \subset \mathcal{A}^{(n)} \subset \cdots \subset \mathcal{A}$ を満たすものとする. また各 n に対して $\mathcal{A}_0^{(n)}$を $\mathcal{A}^{(n)}$の部分σ-加法族とする. このとき, $\{\mathcal{A}_0^{(n)}\}$ が推移列であるとは, 各 n に対して $(\mathcal{A}_0^{(n)}; \mathcal{A}^{(n)}, \mathcal{A}_0^{(n+1)})$ が $\mathcal{P}$に対して Markov となることである. また $\{\mathcal{A}_0^{(n)}\}$ が十分列であるとは, 各 n に対して $\mathcal{A}_0^{(n)}$ suf $(\mathcal{P}; \mathcal{A}^{(n)})$ となることである. 従って, 予測十分性と, 十分性でかつ推移性との関係は, [B54] の結果を用いれば, 次のようになる.

定理 3.1. 次の (i), (ii), (iii) は互いに同値である.

(i) 各 n に対して $\mathcal{A}_0^{(n)}$ pred.suf $(\mathcal{P}; \mathcal{A}^{(n)}, \mathcal{A}_0^{(n+1)})$ である.

(ii) $\{\mathcal{A}_0^{(n)}\}$ が十分列かつ推移列である.

(iii) 各 n に対して $\mathcal{A}_0^{(n)}$ suf $(\mathcal{P}_0^{(n)}; \mathcal{A}^{(n)})$ である.

ここで, 各 n に対して $\mathcal{P}_0^{(n)}$を $dq = gdp$ となる $\mathcal{A}$ 上の確率測度 q全体の集合とする. ただし $p \in \mathcal{P}$で gは非負の $\mathcal{A}_0^{(n+1)}$-可測関数とする.

次に, $X_1, X_2, \ldots$ を確率変数列とし, 各 X_iの標本空間を $(\mathcal{X}, \mathcal{A})$ とし, $\mathcal{P}$を $\mathcal{A}$ 上の確率分布族とする. また, 自然数 $a, b (a \leq b)$ に対して, $\mathcal{A}_a^b$を $(X_a, \ldots, X_b)$ によって誘導されたσ-加法族とする. さらに各 n に対して $\mathcal{B}_1^n$, $\mathcal{C}_1^n$を $\mathcal{A}_1^n$の部分σ-加法族とする. このとき, 定理 2.2 を逐次の場合に適用すれば, 次のようになる.

定理 3.2. 各 n に対して $\mathcal{B}_1^n$ pred.suf $(\mathcal{P}; \mathcal{A}_1^n, \mathcal{A}_{n+1}^{n+1})$ かつ $\mathcal{C}_1^n$ nec $(\mathcal{P}; \mathcal{A}_1^n)$ ならば, 各 n に対して $\mathcal{B}_1^n$ pred.suf $(\mathcal{P}; \mathcal{A}_1^n, \mathcal{C}_1^{n+1})$ である.

証明. 定理 2.2 より, 各 n に対して $(\mathcal{B}_1^n; \mathcal{A}_1^n, \mathcal{B}_1^n \vee \mathcal{A}_{n+1}^{n+1})$ は $\mathcal{P}$に対して Markov であり, $\mathcal{A}_1^n \vee \mathcal{A}_{n+1}^{n+1} = \mathcal{A}_1^{n+1}$であるので, $\mathcal{B}_1^n \vee \mathcal{A}_{n+1}^{n+1}$ suf $(\mathcal{P}; \mathcal{A}_1^{n+1})$ が成り立つ. また, 仮定より $\mathcal{C}_1^{n+1}$ nec $(\mathcal{P}; \mathcal{A}_1^{n+1})$ であるので

$$\mathcal{C}_1^{n+1} \subset \mathcal{B}_1^n \vee \mathcal{A}_{n+1}^{n+1} \ (\mathcal{A}_1^{n+1}, \mathcal{P})$$

が成り立つ. 定理 2.2 の証明と同様にして, $(\mathcal{B}_1^n; \mathcal{A}_1^n, \mathcal{C}_1^{n+1})$ は $\mathcal{P}$に対して Markov となる. 仮定より $\mathcal{B}_1^n$ suf $(\mathcal{P}; \mathcal{A}_1^n)$ は明らかであるので, 各 n に対して $\mathcal{B}_1^n$ pred.suf $(\mathcal{P}; \mathcal{A}_1^n, \mathcal{C}_1^{n+1})$ となる. □

定理 3.2 の命題を統計量を用いて表現すれば, 次のようになる. まず, 確率分布族 $\mathcal{P}$が母数$\theta (\in \Theta)$ によって特徴づけられている, すなわち $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ とする. 次に, $T_n = T_n(X_1, \ldots, X_n)$ が X_{n+1}とθに関して $(X_1, \ldots, X_n)$ に対して予測十分統計量で, U_{n+1}が $X_1, \ldots, X_n, X_{n+1}$に基づく必要統計量, すなわち任意の ($\theta$に関して $(X_1, \ldots, X_n, X_{n+1})$ に対する) 十分統計量 $S_{n+1} = S_{n+1}(X_1, \ldots, X_n, X_{n+1})$ の関数 $U_{n+1} = U_{n+1}(X_1, \ldots, X_n, X_{n+1})$, であるとき, T_nは U_{n+1}とθに関して $(X_1, \ldots, X_n)$ に対して予測十分である.

参考文献

[A82] Akahira, M. (1982). On prediction sufficiency. *Selecta Statistica Canadiana* IV, 1–15.

[AN99] Akahira, M. and Nishihira, Y. (1999). Prediction sufficiency and sequential experimentation. Submitted for publication.

[AT80] Akahira, M. and Takeuchi, K. (1980). A note on prediction sufficiency (adequacy) and sufficiency. *Austral. J. Statist.*, **22**, 332–335.

[B54] Bahadur, R. R. (1954). Sufficiency and statistical decision functions. *Ann. Math. Statist.*, **25**, 423–462.

[L63] Loève, M. (1963). *Probability theory.* (3rd ed.). Van Nostrand, Princeton.

[N99] Nishihira, Y. (1999). Sufficiency and transitivity in sequential decision theory. In press in *İstatistik* Vol. 2, No. 1.

[S67] Skibinsky, M. (1967). Adequate subfields and sufficiency. *Ann. Math. Statist.*, **38**, 155–167.

[SM69] 杉浦誠, 森本治樹 (1969). Adequate σ-field に関する因子分解定理. 数学 **21**, 286–288.

[TA75] Takeuchi, K. and Akahira, M. (1975). Characterizations of prediction sufficiency (adequacy) in terms of risk functions. *Ann. Statist.*, **3**, 1018–1024.

[T77] Torgersen, E. N. (1977). Prediction sufficiency when the loss function does not depend on the unknown parameter. *Ann. Statist.*, **5**, 155–163.