数理解析研究所講究録
第 380 巻 1980 年 44-53

Prediction Sufficiency and Its Relation to Sequential Estimation

赤平 昌文

はじめに

統計的推定論において、条件付期待値によって定義される十分性(sufficiency)の性質および危険十分性(risk sufficiency)すなわち危険関数(risk function)による十分性との関係については、従来研究されている。一方、統計的予測問題においては、いわゆる十分性の概念だけでは不十分であり、もう少し強い意味での十分性の概念が必要になってくる。そこで従来の十分性にMarkov性の条件を付け加えた予測十分性(prediction sufficiency または adequacy)の概念が導入され(Skibinsky [5])、危険十分性による特徴付けが行われた(Takeuchi and Akahira [6])。さらにこの予測十分性の概念を拡張して広義の予測十分性の概念も導入されその特徴付けも研究された([6], Torgersen [7])。ここでは予測十分性の性質について論じ、さらに逐次推定(sequential estimation)の場合にも考察する。

§1. 定義.

$(\mathcal{X}, \mathcal{A})$を標本空間、$\mathcal{A}$上の確率測度の族を$\mathcal{P}$とする。$\mathcal{B}$, $\mathcal{C}$を$\mathcal{A}$の部分σ-代数とし、$\mathcal{B}_0$を$\mathcal{B}$の部分σ-代数とする。集合Aの定義関数をχ_Aで表わす。

定義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、次のことが成り立つことである。これを記号で$\mathcal{B}_0 \ \mathrm{suf}(\mathcal{P}; \mathcal{B})$と表わすことにする。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

$$\phi_B^{\mathcal{B}_0} = E_P(\chi_B \mid \mathcal{B}_0) \quad \text{a.e. } [P]$$

となる。ここで$E_P(\chi_B \mid \mathcal{B}_0)$は、$\mathcal{B}_0$が与えられたときの$\chi_B$のPの下での条件付期待値である。

定義2. $(\mathcal{B}_0; \mathcal{B}, \mathcal{C})$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、

$$E_P(\chi_C \mid \mathcal{B}) = E_P(\chi_C \mid \mathcal{B}_0) \quad \text{a.e. } [P]$$

となることである。

$(\mathcal{B}_0; \mathcal{B}, \mathcal{C})$が$\mathcal{P}$に対してMarkovであるための必要十分条件は、$\mathcal{B}_0$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$B \in \mathcal{B}$と任意の$C \in \mathcal{C}$, 任意の$P \in \mathcal{P}$に対して

$$E_P(\chi_{B \cap C} \mid \mathcal{B}_0) = E_P(\chi_B \mid \mathcal{B}_0)\, E_P(\chi_C \mid \mathcal{B}_0) \quad \text{a.e. } [P]$$

となることである(Loève [4], pp. 351)。

このことから定義2において Markov 性は、$\mathcal{B}$と$\mathcal{C}$について対称である. 定義2において Markov であると呼ぶことについては次のように解釈できる. $X_1, X_2, \cdots$ を $(\mathcal{X}, \mathcal{A}, \mathcal{P})$ 上で定義された確率変数列とし、$\mathcal{B}$, $\mathcal{B}_0$, $\mathcal{C}$ をそれぞれ $(X_1, \cdots, X_n)$, X_n, X_{n+1} によって誘導された σ-代数とすれば、いわゆる Markov 性という意味が理解できる.

定義3. $\mathcal{B}_0$ 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(i), (ii)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.

(i) $\mathcal{B}_0$ *suf* $(\mathcal{P}; \mathcal{B})$;

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

予測十分性を Skibinsky [5] においては adequacy という言葉で呼んでいる.

予測十分性を拡張した概念として広義の予測十分性が、Takeuchi and Akahira [6] において与えられている.

定義4. $\mathcal{B}_0$ が $\mathcal{C}$ と $\mathcal{P}$ に関して $\mathcal{B}$ に対して広義の予測十分であるとは、次の(i'), (ii')が成り立つことである.

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

(ii') $\mathcal{B}_0$-可測集合 B_1 と B_2 が存在して、$P(B_1 \cup B_2)=1$ $(\forall P \in \mathcal{P})$ で、B_1 上では $E_P^{\mathcal{B}}(X.|\mathcal{B}_0)$ は P に無関係で

あり、B_2 上で $E_P^{\mathcal{C}}(\chi.|\mathcal{B}_0)$ は Pに無関係である.

$\mathcal{B}_0$ が$\mathcal{C}$と$\mathcal{P}$に関して$\mathcal{B}$に対して予測十分ならば、広義の予測十分になる. 実際、定義3の(i)は上の(ii′)において、$B_1 = \mathcal{X}$ の場合である.

次に $\mathcal{B}_0$が$\mathcal{C}$と$\mathcal{P}$に関して$\mathcal{B}$に対して広義の予測十分であるが、$\mathcal{B}_0$ pred. suf ($\mathcal{P}$; $\mathcal{B}$, $\mathcal{C}$) とはならない例を示す([6]). $X_1, X_2, \cdots, X_n$ とYが次のような確率変数であると仮定する. $X_1, X_2, \cdots, X_n$ はたがいに独立で同一の正規分布 $N(\theta, 1)$に従っていて、$X_1, X_2, \cdots, X_n$が与えられたときのYの条件付分布は、$\sum_i X_i > a$ または $\sum_i X_i \leqq a$ のときそれぞれ $N(0, 1)$ または $N(\theta, 1)$ である. このとき、$\mathcal{B}$, $\mathcal{B}_0$, $\mathcal{C}$をそれぞれ $(X_1, X_2, \cdots, X_n)$, $\min(a, \sum_i X_i)$, Yによって誘導されたσ-代数とすれば、$\mathcal{B}_0$ は広義の予測十分となるが、$\mathcal{B}_0$は$\mathcal{B}$に対して十分でない.

広義の予測十分性については、Takeuchi and Akahira [6], Torgersen [7] によってその特徴付けが行われており、ここではこの概念についてはこれ以上議論しない.

§2. 予測十分性

予測十分性と十分性の関係については Akahira and Takeuchi [2]によって論じられているが、ここではこれを

さらに発展させた形で考察する（Akahira [1]）.

$\mathcal{A}$ の部分集合族 $\mathcal{A}_1, \mathcal{A}_2$ に対して、$\mathcal{A}_1, \mathcal{A}_2$ を含む最小の σ-代数を $\mathcal{A}_1 \vee \mathcal{A}_2$ によって表わす.

補題 2.1（[2]）. $(\mathcal{X}, \mathcal{A}, P)$ を確率空間とする. $(\mathcal{B}_0; \mathcal{B}, \mathcal{C})$ が P に対して Markov であるための必要十分条件は、任意の $B \in \mathcal{B}$ と任意の $C \in \mathcal{C}$ に対して

$$E_P(\chi_{B \cap C} | \mathcal{B}_0 \vee \mathcal{C}) = \chi_C E_P(\chi_B | \mathcal{B}_0) \quad a.e. [P]$$

となることである.

この補題はこの節の定理を証明するために非常に有効である.

定理 2.1（[2]）. $\mathcal{S}$ を $\mathcal{B}_0 \vee \mathcal{C}$ の部分 σ-代数とする. $\mathcal{B}_0$ pred. suf $(\mathcal{P}; \mathcal{B}, \mathcal{C})$ でかつ $\mathcal{S}$ 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.

次の定理 2.2, 2.3, 2.4 のために、$\mathcal{P}$ は homogeneous である、すなわち任意の $P, Q \in \mathcal{P}$ に対して P と Q がたがいに絶対連続であると仮定する.

定理 2.1 の命題の逆は、次の定理 2.2 のような形で成り立つ.

定理 2.2（[2]）. $(\mathcal{B}_0; \mathcal{B}, \mathcal{C})$ は $\mathcal{P}$ に対して Markov でありかつ $\mathcal{B}_0 \vee \mathcal{C}$ suf $(\mathcal{P}; \mathcal{B} \vee \mathcal{C})$ ならば $\mathcal{B}_0$ suf $(\mathcal{P};$

$\mathcal{B}$)である．

定理2.2において $\mathcal{P}$ が homogeneous であるという仮定が必要であることが次の反例によって示される．

<u>例2.1.</u> $\mathcal{B}=\mathcal{C}=\{\phi, B, B^c, \mathcal{X}\}$, $\mathcal{B}_0=\{\phi, \mathcal{X}\}$ とし、$\mathcal{P}=\{P, Q\}$ で $P(B)=Q(B^c)=1$ とする．このとき $(\mathcal{B}_0;\mathcal{B},\mathcal{C})$ は $\mathcal{P}$ に対して Markov であり、$\mathcal{B}_0\vee\mathcal{C}$ suf $(\mathcal{P};\mathcal{B}\vee\mathcal{C})$ であるが、$\mathcal{B}_0$ は $\mathcal{P}$ に関して $\mathcal{B}$ に対して十分でない．

次の例によって、定理2.2において Markov 性の仮定が必要であることを示す．

<u>例2.2</u> ([2]). (X, Y, Z) は平均 $(\theta, 0, \theta)$ と後に定義するような既知の分散、共分散をもつ同時正規分布に従っていると仮定する．ただしパラメータ θ は未知とする．$\mathcal{B}_0, \mathcal{B}, \mathcal{C}$ をそれぞれ X, (X, Y), Z によって誘導された σ-代数とする．$Y=X-Z+U$ であると仮定する．ここで U は X, Z と独立とする．さらに $Cov(X, Z)=0$ で $Var(X)=Var(Z)=Var(U)=1$ であることを仮定する．このとき $E(U)=0$, $Var(Y)=3$, $Cov(Y, X)=1$, $Cov(Y, Z)=-1$ となるから、$\mathcal{B}_0\vee\mathcal{C}$ は $\mathcal{B}\vee\mathcal{C}$ に対して十分であるが、$\mathcal{B}_0$ は $\mathcal{B}$ に対して十分ではなく、また $(\mathcal{B}_0;\mathcal{B},\mathcal{C})$ は Markov でないことが示される．

$\mathcal{D}$を$\mathcal{A}$の部分σ-代数、$\mathcal{C}_0$を$\mathcal{C}$の部分σ-代数とする.

<u>定理 2.3</u> ([1]).

$\mathcal{B}_0$ pred. suf $(\mathcal{P}; \mathcal{B}, \mathcal{C}\vee\mathcal{D})$であるための必要十分条件は、$(\mathcal{B}_0; \mathcal{B}, \mathcal{C})$が$\mathcal{P}$に対してMarkovでかつ$\mathcal{B}_0\vee\mathcal{C}$ pred. suf $(\mathcal{P}; \mathcal{B}\vee\mathcal{C}, \mathcal{D})$となることである.

定理 2.3 の証明において、$\mathcal{P}$が homogeneous であるという仮定は十分性の証明にのみ必要である.

定理 2.3 で $\mathcal{D}=\mathcal{B}_0\vee\mathcal{C}$ とすれば、定理 2.2 の結果が出てくることがわかる.

<u>定理 2.4</u> ([1]).

次の条件 (i), (ii), (iii) について考える.

(i) $\mathcal{B}_0$ pred. suf $(\mathcal{P}; \mathcal{B}, \mathcal{C}\vee\mathcal{D})$ でかつ $\mathcal{C}_0$ pred. suf $(\mathcal{P}; \mathcal{C}, \mathcal{B}\vee\mathcal{D})$である.

(ii) $(\mathcal{B}_0; \mathcal{B}, \mathcal{C})$と$(\mathcal{C}_0; \mathcal{B}, \mathcal{C})$は$\mathcal{P}$に対してMarkovであって、$\mathcal{B}_0\vee\mathcal{C}$ pred. suf $(\mathcal{P}; \mathcal{B}\vee\mathcal{C}, \mathcal{D})$でかつ $\mathcal{B}\vee\mathcal{C}_0$ pred. suf $(\mathcal{P}; \mathcal{B}\vee\mathcal{C}, \mathcal{D})$である.

(iii) $\mathcal{B}_0\vee\mathcal{C}_0$ pred. suf $(\mathcal{P}; \mathcal{B}\vee\mathcal{C}, \mathcal{D})$.

このとき (i) と (ii) は同値であり、(i) (または (ii)) ならば (iii) である.

「(i) $\Leftrightarrow$ (ii)」の証明は定理 2.3 から明らかである. $\mathcal{P}$が homogeneous であるという仮定は、「(ii) $\Rightarrow$ (i)」の

証明にのみ必要である.

§3. 逐次推定における予測十分性

Bahadur [3] によって transitivity という概念が導入され、十分性でかつ transitivity の特徴付けが論じられている. $\{\mathcal{A}^{(n)}\}$ を $\mathcal{A}$ の部分 σ-代数列で $\mathcal{A}^{(1)} \subset \mathcal{A}^{(2)} \subset \cdots \subset \mathcal{A}^{(n)} \subset \cdots \subset \mathcal{A}$ 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-代数とする. $\{\mathcal{A}_0^{(n)}\}$ が transitiveな列であるとは 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 $\{\mathcal{A}_0^{(n)}\}$ が十分な列であるとは、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 transitivity との関係は、Bahadur [3] によって次の定理のようになる.

定理 3.1.

次の条件(i), (ii), (iii) はたがいに同値である.

(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 十分でかつ transitive な列である.

(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 $dQ = g\,dP$ 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 は非負で $\mathcal{A}_0^{(n+1)}$-可測関数とする．

「(i)$\Leftrightarrow$(ii)」の証明は定義から明らかである．「(ii)$\Leftrightarrow$(iii)」は Bahadur [3] の定理 11.3 である．Skibinsky [5] の定理1の「(i)$\Leftrightarrow$(iv)」は、定理 3.1 の「(i)$\Leftrightarrow$(iii)」（すなわち Bahadur [3] の定理 11.3）から直接出てくることがわかる．

$X_1, X_2, \cdots$ を $(\mathcal{X}, \mathcal{A}, \mathcal{P})$ 上で定義された確率変数列とする．$a\leq b$ に対して $\mathcal{A}_a^b$ を $(X_a, \cdots, X_b)$ によって誘導された σ-代数とする．各 n に対して $\mathcal{B}_1^n$ を $\mathcal{A}_1^n$ の部分 σ-代数とする．ここで定理 3.2 のために、次のことを仮定する．各 n と任意の $B\in\mathcal{B}_1^{n+1}$ に対して $\mathcal{B}_1^n\vee\mathcal{A}_{n+1}^{n+1}$-可測関数 φ_B^{n+1} が存在して、任意の $P\in\mathcal{P}$ に対して

$$E_P(\varphi_B^{n+1}|\mathcal{A}_1^n) = E_P(\chi_B|\mathcal{A}_1^n) \quad \text{a.e. } [P]\ ;$$
$$E_P(\varphi_B^{n+1}|\mathcal{B}_1^n) = E_P(\chi_B|\mathcal{B}_1^n) \quad \text{a.e. } [P]$$

となる．

<u>定理 3.2.</u> 各 n に対して $\mathcal{B}_1^n$ pred. suf $(\mathcal{P}; \mathcal{A}_1^n, \mathcal{A}_{n+1}^{n+1})$ であるならば、各 n に対して $\mathcal{B}_1^n$ pred. suf $(\mathcal{P}; \mathcal{A}_1^n, \mathcal{B}_1^{n+1})$ である．

証明の概略．

各 n について、$(\mathcal{B}_1^n; \mathcal{A}_1^n, \mathcal{A}_{n+1}^{n+1})$ が $\mathcal{P}$ に対して Markov で

あると仮定する．このとき各n，任意の$B \in \mathcal{B}_1^{n+1}$と任意の$P \in \mathcal{P}$に対して

$$E_P(\chi_B | \mathcal{A}_1^n) = E_P(\varphi_B^{n+1} | \mathcal{A}_1^n) = E_P(\varphi_B^{n+1} | \mathcal{B}_1^n)$$
$$= E_P(\chi_B | \mathcal{B}_1^n) \qquad a.e.\ [P]$$

であるから、各nについて$(\mathcal{B}_1^n ; \mathcal{A}_1^n, \mathcal{B}_1^{n+1})$は$\mathcal{P}$に対して Markov となる．

References

[1] Akahira,M.: On prediction sufficiency. (to appear).

[2] Akahira,M. and Takeuchi,K.: A note on prediction sufficiency (adequacy) and sufficiency. To be published in the Austral.J.Statist., 22 (1980).

[3] Bahadur,R.R.: Sufficiency and statistical decision functions. Ann.Math.Statist., 25 (1954), 423-462.

[4] Loève,M.: Probability Theory. (3rd ed.). Van Nostrand, Princeton (1963).

[5] Skibinsky,M.: Adequate subfields and sufficiency. Ann.Math. Statist., 38 (1967), 155-161.

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

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