

On approximate Sufficiency

T. Kusama (Waseda Univ-)

H. Kudo defined the notion of approximate sufficiency as follows [1]. Let $(X, \mathcal{A}, \{P, Q\})$ be a statistical structure and $\{\mathcal{A}_n\}$ be a sequence of sub- σ -algebra such that $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \dots$ and $\bigvee_{n=1}^{\infty} \mathcal{A}_n = \mathcal{A}$. A sequence $\{\mathcal{B}_n\}$ ($\mathcal{B}_n \subset \mathcal{A}_n$) is called approximately sufficient for $\{P, Q\}$ if there exist probability measures P_n, Q_n ($n=1, 2, \dots$) on $\mathcal{B}_n$ satisfying

$$(1) \|P - P_n\|_{\mathcal{A}_n} \rightarrow 0, \|Q - Q_n\|_{\mathcal{A}_n} \rightarrow 0 \quad (n \rightarrow \infty)$$

$$(2) \mathcal{B}_n \text{ is sufficient for } \{P_n, Q_n\}.$$

Here $\|P - P_n\|_{\mathcal{A}_n}$ means $\sup_{B \in \mathcal{A}_n} |P(B) - P_n(B)|$.

In [3] the author extends the notion of approximate sufficiency to general statistical structures. Let $(X, \mathcal{A}, \mathcal{P} = \{P_\theta \mid \theta \in \Omega\})$ be a statistical structure. Let $\{\mathcal{A}_n\}$ be the same as defined above. A sequence $\{\mathcal{B}_n\}$ ($\mathcal{B}_n \subset \mathcal{A}_n$) is called approximately sufficient for $\mathcal{P}$ if there exist families of probability measures $\mathcal{P}_n = \{P_{\theta,n} \mid \theta \in \Omega\}$ on $\mathcal{B}_n$ ($n=1, 2, \dots$) satisfying

- (1) $\|P_\theta - P_{\theta, n}\|_{\mathcal{O}_n} \rightarrow 0 \quad (n \rightarrow \infty) \quad (\forall \theta \in \Omega)$
 (2) $\mathcal{B}_n$ is sufficient for $\mathcal{P}_n \quad (n=1, 2, \dots)$.

Then we have results corresponding to the well-known results about sufficiency obtained by Halmos-Savage and Bahadur [4] [5].

Theorem 1 Let $\mathcal{O}$ be σ -generated and $\mathcal{P}$ -be dominated. Then, by choosing a dominating probability measure λ_0 suitably, the following four assertions are equivalent [3].

- (a) $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}$
 (b) $\tilde{P}_{\lambda_0}(f_\theta, L_{\lambda_0}(\mathcal{B}_n)) \rightarrow 0 \quad (n \rightarrow \infty) \quad (\forall \theta \in \Omega)$
 (c) $P_{\lambda_0}(f_\theta, E_{\lambda_0}(f_\theta | \mathcal{B}_n)) \rightarrow 0 \quad (n \rightarrow \infty) \quad (\forall \theta \in \Omega)$
 (d) $\mathcal{B}_0 = \lambda_0\text{-}\liminf \mathcal{B}_n$ is sufficient for $\mathcal{P}$.

Here P_{λ_0} denotes the distance in $L^1(X, \mathcal{O}, \lambda_0)$ and $L_{\lambda_0}(\mathcal{B}_n)$ denotes the set of all $\mathcal{B}_n$ -measurable elements in $L^1(X, \mathcal{O}, \lambda_0)$. $\tilde{P}_{\lambda_0}(f_\theta, L_{\lambda_0}(\mathcal{B}_n))$ denotes the distance between $f_\theta = \frac{dP_\theta}{d\lambda_0}$ and the set $L_{\lambda_0}(\mathcal{B}_n)$ in $L^1(X, \mathcal{O}, \lambda_0)$ and $\lambda_0\text{-}\liminf \mathcal{B}_n$ denotes $\{A \in \mathcal{O} \mid \exists B_n \in \mathcal{B}_n : \lambda_0(A \Delta B_n) \rightarrow 0\}$. $\lambda_0\text{-}\liminf \mathcal{B}_n$ is called the lower limit of $\{\mathcal{B}_n\}$. It is proved in [2] that $\lambda_0\text{-}\liminf \mathcal{B}_n$ is a σ -algebra.

In [2] λ_0 - $\liminf \mathcal{B}_n$ is characterized as the σ -algebra $\mathcal{B}_0$ having the following properties.

(i) $\mathcal{B}_0$ satisfies

$$\liminf_{n \rightarrow \infty} \int |E_{\lambda_0}(f | \mathcal{B}_n)| d\lambda_0 \geq \int |E_{\lambda_0}(f | \mathcal{B}_0)| d\lambda_0$$

for every bounded $\mathcal{A}$ -measurable f [A]

(ii) any σ -algebra $\mathcal{B}$ satisfying [A] is contained in $\mathcal{B}_0$.

If, for every $\theta_1, \theta_2 \in \Omega$, $\{\mathcal{B}_n\}$ is approximately sufficient for $\{P_{\theta_1}, P_{\theta_2}\}$, $\{\mathcal{B}_n\}$ is said to be pairwise approximately sufficient for $\mathcal{P}$.

Theorem 2 Under the same assumptions as those in the above theorem, approximate sufficiency and pairwise approximate sufficiency are equivalent.

Theorem 3 If $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}$, for every $\mathcal{A}$ -measurable, bounded f , there exist a sequence $\{h_n\}$ of $\mathcal{A}$ -measurable, bounded functions and versions $\tilde{E}_{\theta}(f | \mathcal{B}_n)$ of $E_{\theta}(f | \mathcal{B}_n)$ such that $P_{\lambda_0}(\tilde{E}_{\theta}(f | \mathcal{B}_n), h_n) \rightarrow 0$ ($n \rightarrow \infty$) for every $\theta \in \Omega$.

The converse of this theorem is an open

problem. But, when $\{\mathcal{B}_n\}$ is monotone increasing, the converse holds. The question naturally arises whether the diameter of $\{\tilde{E}_\theta(f|\mathcal{B}_n) \mid \theta \in \Omega\}$ tends to 0 as $n \rightarrow \infty$ by choosing suitable versions $\tilde{E}_\theta(f|\mathcal{B}_n)$. The answer to this question is generally negative. But, if Ω is compact with respect to the metric $d(\theta_1, \theta_2) = \sup_{B \in \mathcal{C}} |P_{\theta_1}(B) - P_{\theta_2}(B)|$ and $\mathcal{P}$ is homogeneous, the answer is positive.

Let $\mathcal{C}$ be a σ -algebra satisfying $\mathcal{B} \subset \mathcal{C} \subset \mathcal{A}$. If, for every $\mathcal{C}$ -measurable f , there exists a conditional expectation $E(f|\mathcal{B})$ common to every P_θ , $\mathcal{B}$ is said to be $\mathcal{C}$ -sufficient for $\mathcal{P}$.

Theorem 4 If there exists a sequence of σ -algebras $\{\mathcal{C}_n\}$ such that $\mathcal{B}_n \subset \mathcal{C}_n \subset \mathcal{A}$, λ_0 - $\liminf \mathcal{C}_n = \mathcal{A}$ and $\mathcal{B}_n$ is $\mathcal{C}_n$ -sufficient for $\mathcal{P}$, then $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}$.

The converse of this theorem is an open problem.

~~~~~

[1] H. Kudō : On an approximation to a sufficient statistics including a concept of asymptotic sufficiency, J. Fac. Sci., Univ of

Tokyo, Sec I. 17 (1970), pp 273-290

[2] H. Kudō : A note on the strong convergence of  $\sigma$ -algebras; Ann. Probability 2 (1974) pp 76~83.

[3] T. Kusama; On approximate sufficiency;

Osaka Journal of Mathematics Vol 13, No 3, 1976, pp 661-669

[4] P. R. Halmos and J. L. Savage : Application of the Radon-Nikodym theorem to the theory of sufficient statistics, Ann. Math. Statist 20 (1949) pp 225~241.

[5] R. R. Bahadur; Sufficiency and statistical decision functions, Ann. Math. Statist 25 (1954) pp 423~462.