Kusama, T.
Osaka J. Math.
**13** (1976), 661–669

# ON APPROXIMATE SUFFICIENCY

TOKITAKE KUSAMA

(Received October 10, 1975)
(Revised March 10, 1976)

H. Kudō defined the notion of approximate sufficiency in his paper ([4], [6]) and proved some interesting results. In this paper we obtain some characterizations for it.

## 1. Notations and definitions

Let $(X, \mathcal{A})$ be a sample space consisting of a set $X$ and a $\sigma$-algebra $\mathcal{A}$ of subsets of $X$. The reader should understand by the word "$\sigma$-algebra" and "algebra" a sub-$\sigma$-algebra and subalgebra of $\mathcal{A}$, respectively. Given a $\sigma$-algebra $\mathcal{B}$ and a finite measure $\lambda$ on $\mathcal{A}$, $E_\lambda(f\,|\,\mathcal{B})$ denotes the conditional expectation of a $\lambda$-integrable function $f$ over $X$ given $\mathcal{B}$ with respect to $\lambda$: *i.e.*, $E_\lambda(f\,|\,\mathcal{B})$ is a $\mathcal{B}$-measurable function such that $\int_B f d\lambda = \int_B E_\lambda(f\,|\,\mathcal{B}) d\lambda$ for every $B \in \mathcal{B}$. When a probability measure $P$ on $\mathcal{A}$ is absolutely continuous with respect to $\lambda$ (we write $P \ll \lambda$), $\frac{dP}{d\lambda}$ denotes the Radon-Nikodym derivative. It is clear that $E_\lambda\left(\frac{dP}{d\lambda}\,|\,\mathcal{B}\right)$ coincides with the Radon-Nikodym derivative $\left.\frac{dP}{d\lambda}\right|_{\mathcal{B}}$ of $P/\mathcal{B}$ with respect to $\lambda/\mathcal{B}$, where $P/\mathcal{B}$ and $\lambda/\mathcal{B}$ are the contractions of $P$ and $\lambda$ to $\mathcal{B}$ respectively.

For a finite signed measure $m$, $||m||_{\mathcal{B}}$ denotes the value $\sup_{B\in\mathcal{B}} |m(B)|$. When $m \ll \lambda$ and $m(X)=0$, it is well known that $||m||_{\mathcal{B}} = \frac{1}{2}\int_X \left|\left.\frac{dm}{d\lambda}\right|_{\mathcal{B}}\right| d\lambda$ $\left(= \frac{1}{2}\int_X \left|E_\lambda\left(\frac{dm}{d\lambda}\,|\,\mathcal{B}\right)\right| d\lambda\right)$. Here and hereafter the integration without any assignment of its domain should be understood as that extended over the whole space $X$.

Let $\{\mathcal{A}_n\}$ be an increasing sequence of $\sigma$-algebras and $\{\mathcal{B}_n\}$ a sequence of $\sigma$-algebras satisfying $\mathcal{B}_n \subset \mathcal{A}_n$. According to Kudō ([4], [6]), $\{\mathcal{B}_n\}$ is said to be approximately sufficient for a pair $\{P, Q\}$ of probability measures on $\mathcal{A}$, if for each $n$ there is a pair of probability measures $\{P_n, Q_n\}$ on $\mathcal{A}_n$ such that

$\lim_{n\to\infty}||P_n-P||_{\mathcal{A}_n}=\lim_{n\to\infty}||Q_n-Q||_{\mathcal{A}_n}$ and that $\mathcal{B}_n$ is sufficient for $\{P_n, Q_n\}$ on $\mathcal{A}_n$ for every $n$. We shall consider this notion in the case of an arbitrary family of probability measures.

REMARK. A slight errata in Kudō's definition of approximate sufficiency in [4] is corrected in [6].

Let $\mathcal{P}=\{P_\theta|\theta\in\Omega\}$ be a family of probabuility measures defined on $\mathcal{A}$, where $\Omega$ is a parameter space. A sequence $\{\mathcal{B}_n\}$ of $\sigma$-algebras is said to be approximately sufficient for $\mathcal{P}$ if for each $n$ there is a family of probability measures $\mathcal{P}_n=\{P_{\theta,n}|\theta\in\Omega\}$ on $\mathcal{A}_n$ such that $\lim_{n\to\infty}||P_{\theta,n}-P_\theta||_{\mathcal{A}_n}=0$ for all $\theta\in\Omega$ and that $\mathcal{B}_n$ is sufficient for $\mathcal{P}_n$ on $\mathcal{A}_n$ for every $n$. Throughout this paper we assume

(A1) $\bigvee_{n=1}^{\infty}\mathcal{A}_n=\mathcal{A}$,

where $\bigvee_{n=1}^{\infty}\mathcal{A}_n$ denotes the $\sigma$-algebra generated by $\{\mathcal{A}_n\}$, and assume that

(A2) $\mathcal{P}$ is dominated by a finite measure $\lambda$ on $\mathcal{A}$.

(A3) $\mathcal{A}$ is countably generated.

Let $L^1(X, \mathcal{A}, \lambda)$ be the space of all $\lambda$-integrable, real valued, $\mathcal{A}$-measurable functions defined on $X$ with the metric $\rho_\lambda(f, g)=\int|f-g|d\lambda$. The distance between $f\,(\in L^1(X, \mathcal{A}, \lambda))$ and $A\,(\subset L^1(X, \mathcal{A}, \lambda))$ is defined by $\tilde{\rho}_\lambda(f, A)=\inf_{g\in A}\rho_\lambda(f, g)$. Let $L_\lambda(\mathcal{B})$ denote the set of all $\mathcal{B}$-measurable elements, which is a subspace of $L^1(X, \mathcal{A}, \lambda)$.

Let $\{\mathcal{B}_n\}$ be a sequence of $\sigma$-algebras. The subfamily of $\mathcal{A}$ consisting of $B\,(\in\mathcal{A})$ for which there are $B_n\in\mathcal{B}_n$ such that $\lambda(B\triangle B_n)\to 0$ $(n\to\infty)$ is called the lower limit of $\{\mathcal{B}_n\}$ and denoted as $\lambda$-liminf $\mathcal{B}_n$. Here $B\triangle B_n$ means symmetric difference of $B$ and $B_n$. $\lambda$-liminf $\mathcal{B}_n$ is a $\sigma$-algebra ([5] Theorem 3.2).

Since $\mathcal{P}$ is domianted and $\mathcal{A}$ is countably generated, there exists $\Omega^*=\{\theta_1, \theta_2, \cdots\}$ of $\Omega$ such that $\mathcal{P}^*=\{P_\theta|\theta\in\Omega^*\}$ is dense in $\mathcal{P}$ ([1]). Let $\lambda_0=\sum_{i=1}^{\infty}\beta_i P_{\theta_i}$ $(\beta_i>0, \sum_{i=1}^{\infty}\beta_i<\infty)$. Then it is easy to see that $\lambda_0$ is equivalent to $\mathcal{P}$ (we write $\lambda_0\approx\mathcal{P}$). We write $f_\theta=\frac{dP_\theta}{d\lambda_0}$.

## 2. Some characterizations for approximate sufficiency

**Theorem 1.** *Under Assumptions* $(A1)\sim(A3)$ *in* §1, *the following four assertions are all equivalent.*

(a) $\{\mathcal{B}_n\}$ *is approximately sufficient for* $\mathcal{P}$.

(b) $\tilde{\rho}_{\lambda_0}(f_\theta, L_{\lambda_0}(\mathcal{B}_n))\to 0$ $(n\to\infty)$ *for every* $\theta\in\Omega$.

(c) $\rho_{\lambda_0}(f_\theta, E_{\lambda_0}(f_\theta | \mathcal{B}_n)) \to 0$ $(n \to \infty)$ *for every* $\theta \in \Omega$.
(d) $\mathcal{B}_0 = \lambda_0$-*liminf* $\mathcal{B}_n$ *is sufficient for* $\mathcal{P}$.

Proof. (a)$\Rightarrow$(b). Since $\bigvee_{n=1}^{\infty} \mathcal{A}_n = \mathcal{A}$ by assumption in §1 we have

$$(1) \quad \rho_{\lambda_0}(E_{\lambda_0}(f_\theta | \mathcal{A}_n), f_\theta) = \rho_{\lambda_0}(E_{\lambda_0}(f_\theta | \mathcal{A}_n), E_{\lambda_0}(f_\theta | \mathcal{A})) \to 0 \ (n \to \infty)$$

for every $\theta \in \Omega$ ([7]).

Since $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}$, there exist $\mathcal{P}_n = \{P_{\theta,n} | \theta \in \Omega\}$ $(n=1, 2, \cdots)$ on $\mathcal{A}_n$ such that $\lim_{n \to \infty} ||P_{\theta,n} - P_\theta||_{\mathcal{A}_n} = 0$ for every $\theta \in \Omega$ and that $\mathcal{B}_n$ is sufficient for $\mathcal{P}_n$. Define $\lambda_n = \sum_{i=1}^{\infty} \beta_i P_{\theta_i,n}$ on $\mathcal{A}_n$ with $\theta_i \in \Omega^*$. Hence we have

$$(2) \qquad ||\lambda_n - \lambda_0||_{\mathcal{A}_n} \to 0 \qquad (n \to \infty).$$

Putting $f_{\theta_i,n} = \frac{dP_{\theta_i,n}}{d\lambda_n}$ for every $i$, we have

$$(3) \quad \begin{aligned} ||f_{\theta_i,n} d\lambda_0 - P_{\theta_i}||_{\mathcal{A}_n} &\leqq ||f_{\theta_i,n} d\lambda_0 - f_{\theta_i,n} d\lambda_n||_{\mathcal{A}_n} + ||f_{\theta_i,n} d\lambda_n - P_{\theta_i}||_{\mathcal{A}_n} \\ &= ||f_{\theta_i,n} d\lambda_0 - f_{\theta_i,n} d\lambda_n||_{\mathcal{A}_n} + ||P_{\theta_i,n} - P_{\theta_i}||_{\mathcal{A}_n}. \end{aligned}$$

The first term of the right hand side of (3) tends to 0 as $n \to \infty$ from $f_{\theta_i,n}(x) \leqq \beta_i^{-1}$ for all $x$ and (2) and by assumption the second term tends also to 0 as $n \to \infty$. So we have

$$(4) \qquad ||f_{\theta_i,n} d\lambda_0 - P_{\theta_i}||_{\mathcal{A}_n} \to 0 \qquad (n \to \infty)$$

for every $i$. It follows from (1), (4) and the $\mathcal{A}_n$-maesurability of $f_{\theta_i,n}$ that

$$(5) \quad \begin{aligned} \rho_{\lambda_0}(f_{\theta_i,n}, f_{\theta_i}) &\leqq \rho_{\lambda_0}(f_{\theta_i,n}, E_{\lambda_0}(f_{\theta_i} | \mathcal{A}_n)) + \rho_{\lambda_0}(E_{\lambda_0}(f_{\theta_i} | \mathcal{A}_n), f_{\theta_i}) \\ &= \int \Big| f_{\theta_i,n} - \frac{dP_{\theta_i}}{d\lambda_0}\Big|_{\mathcal{A}_n} \Big| d\lambda_0 + \rho_{\lambda_0}(E_{\lambda_0}(f_{\theta_i} | \mathcal{A}_n), f_{\theta_i}) \\ &\leqq 2||f_{\theta_i,n} d\lambda_0 - P_{\theta_i}||_{\mathcal{A}_n} + \rho_{\lambda_0}(E_{\lambda_0}(f_{\theta_i} | \mathcal{A}_n), f_{\theta_i}) \\ &\to 0 \qquad (n \to \infty) \end{aligned}$$

for every $i$. $f_{\theta_i,n}$ is not only $\mathcal{A}_n$-measurable but also $\mathcal{B}_n$-measurable since $\mathcal{B}_n$ is sufficient for $\mathcal{P}_n$ ([3] Theorem 1). The $\mathcal{B}_n$-measurability of $f_{\theta_i,n}$ and (5) imply

$$(6) \qquad \tilde{\rho}_{\lambda_0}(f_{\theta_i}, L_{\lambda_0}(\mathcal{B}_n)) \to 0 \qquad (n \to \infty).$$

As $\mathcal{P}^*$ is dense in $\mathcal{P}$, it follows from (6) that $\tilde{\rho}_{\lambda_0}(f_\theta, L_{\lambda_0}(\mathcal{B})) \to 0$ $(n \to \infty)$ for every $\theta \in \Omega$.

(b)$\Rightarrow$(c). Since $\tilde{\rho}_{\lambda_0}(f_\theta, L_{\lambda_0}(\mathcal{B}_n)) \to 0$ $(n \to \infty)$ by assumption, there exist $\mathcal{B}_n$-measurable $g_{\theta,n} \geqq 0$ $(n=1, 2, \cdots; \theta \in \Omega)$ such that

$$(7) \qquad \rho_{\lambda_0}(f_\theta, g_{\theta,n}) \to 0 \qquad (n \to \infty)$$

for every $\theta\in\Omega$. Since $g_{\theta,n}$ and $E_{\lambda_0}(f_\theta|\mathcal{B}_n)$ are $\mathcal{B}_n$-measurable, we have by (7)

$$(8)\qquad \begin{aligned}\rho_{\lambda_0}(g_{\theta,n}, E_{\lambda_0}(f_\theta|\mathcal{B}_n)) &\leqq 2||g_{\theta,n}d\lambda_0-E_{\lambda_0}(f_\theta|\mathcal{B}_n)d\lambda_0||_{\mathcal{B}_n}\\ &=2||g_{\theta,n}d\lambda_0-f_\theta d\lambda_0||_{\mathcal{B}_n}\\ &\leqq 2||g_{\theta,n}d\lambda_0-f_\theta d\lambda_0||_{\mathcal{A}}\\ &\leqq 2\rho_{\lambda_0}(g_{\theta,n}, f_\theta)\to 0 \qquad (n\to\infty)\end{aligned}$$

for every $\theta\in\Omega$. It follows from (7), (8) that

$$\rho_{\lambda_0}(f_\theta, E_{\lambda_0}(f_\theta|\mathcal{B}_n))\leqq\rho_{\lambda_0}(f_\theta, g_{\theta,n})+\rho_{\lambda_0}(g_{\theta,n}, E_{\lambda_0}(f_\theta|\mathcal{B}_n))\to 0 \qquad (n\to\infty)$$

for every $\theta\in\Omega$. This establishes (c).

(c)$\Rightarrow$(d). It suffices to prove the $\mathcal{B}_0$-measurability of $f_\theta$ ([13] Theorem 1). For this purpose it is sufficient to prove $\{f_\theta\geqq a\}\in\mathcal{B}_0$ for a real number a in a dense set $A$ of the real line. Since $A_\theta=\{a|\lambda_0(\{f_\theta=a\})=0\}$ is dense, we shall prove $\{f_\theta\geqq a\}\in\mathcal{B}_0$ for $a\in A_\theta$. Writing $g_{\theta,n}=E_{\lambda_0}(f_\theta|\mathcal{B}_n)$, we have $\rho_{\lambda_0}(f_\theta, g_{\theta,n})\to 0$ $(n\to\infty)$ by assumption. We prove $\lambda_0(\{f_\theta\geqq a\}\triangle\{g_{\theta,n}\geqq a\})\to 0$ $(n\to\infty)$ for $a\in A_\theta$. Let $\varepsilon$ be a given positive number. Then

$$\begin{aligned}\lambda_0(\{f_\theta\geqq a\}\triangle\{g_{\theta,n}\geqq a\})&=\lambda_0(\{f_\theta\geqq a, g_{\theta,n}<a\})+\lambda_0(\{f_\theta<a, g_{\theta,n}\geqq a\})\\ &\leqq\lambda_0(\{f_\theta\geqq a+\varepsilon, g_{\theta,n}<a\})+\lambda_0(\{a\leqq f_\theta<a+\varepsilon\})\\ &\quad+\lambda_0(\{f_\theta<a-\varepsilon, g_{\theta,n}\geqq a\})+\lambda_0(\{a-\varepsilon\leqq f_\theta<a\})\\ &\leqq\lambda_0(\{|g_{\theta,n}-f_\theta|>\varepsilon\})+\lambda_0(\{|f_\theta-a|\leqq\varepsilon\})\\ &\quad\to\lambda_0(\{|f_\theta-a|\leqq\varepsilon\}) \quad\text{as}\quad n\to\infty\,.\end{aligned}$$

Since $\varepsilon$ is arbitrary and $\lambda_0(\{f_\theta=a\})=0$ by assumption, we have $\lim_{n\to\infty}\lambda_0(\{f_\theta\geqq a\}\triangle\{g_{\theta,n}\geqq a\})=0$. From $\{g_{\theta,n}\geqq a\}\in\mathcal{B}_n$ and the definition of $\mathcal{B}_0$, it follows that $\{f_\theta\geqq a\}\in\mathcal{B}_0$.

(d)$\Rightarrow$(a). At first we shall prove that for a given $\varepsilon>0$, there exist $n_0$ and non-negative $\mathcal{B}_n$-measurable $g_{\theta,n}$ for $n\geqq n_0$ such that $\rho_{\lambda_0}(f_\theta, g_{\theta,n})<\varepsilon$ and $E_{\lambda_0}(g_{\theta,n})>0$. Perhaps $n_0$ may depend on $\theta$. Since $\mathcal{B}_0$ is sufficient by assumption, $f_\theta$ is $\mathcal{B}_0$-measurable. Hence there exists a non-negative $\mathcal{B}_0$-measurable function $h_\theta=\sum_{i=1}^{k_\theta}\alpha_{\theta,i}I_{A_{\theta,i}}$ with $\mathcal{B}_0$-measurable sets $A_{\theta,i}$ such that $\rho_{\lambda_0}(f_\theta, h_\theta)<\frac{\varepsilon}{2}$, where $I_A$ is the defining function of $A$. Consequently there exists an $n_0$ such that for each $n\geqq n_0$ we can choose $C_{n,1}, C_{n,2}, \cdots, C_{n,k_\theta}$ from $\mathcal{B}_n$ satisfying $\lambda_0(A_{\theta,i}\triangle C_{n,i})<\frac{\varepsilon}{2k_\theta\max(\alpha_{\theta,1}, \cdots, \alpha_{\theta,k_\theta})}$. We note that $n_0$, $C_{n,i}$ may depend on $\theta$. $g_{\theta,n}=\sum_{i=1}^{k_\theta}\alpha_{\theta,i}I_{C_{n,i}}$ is $\mathcal{B}_n$-measurable and we have for $n\geqq n_0$

$$(9)\qquad \rho_{\lambda_0}(h_\theta, g_{\theta,n}) \leqq \sum_{i=1}^{k\theta} \alpha_{\theta,i} \int |I_{A_{\theta,i}} - I_{C_{n,i}}| d\lambda_0 = \sum_{i=1}^{k\theta} \alpha_{\theta,i} \lambda_0(A_{\theta,i} \triangle C_{n,i}) < \frac{\varepsilon}{2}.$$

$\rho_{\lambda_0}(f_\theta, h_\theta) < \frac{\varepsilon}{2}$ and (9) yield $\rho_{\lambda_0}(f_\theta, g_{\theta,n}) < \varepsilon$ for $n \geqq n_0$. Thus we have proved that, for a given $\varepsilon > 0$, there exist $n_0$ and $\mathcal{B}_n$-measurable $g_{\theta,n}$ for $n \geqq n_0$ such that $g_{\theta,n} \geqq 0$, $E_{\lambda_0}(g_{\theta,n}) > 0$ and $\rho_{\lambda_0}(f_\theta, g_{\theta,n}) < \varepsilon$.

Let a $\mathcal{B}_n$-measurable $h_{\theta,n}$ be such that $h_{\theta,n} \geqq 0$, $E_{\lambda_0}(h_{\theta,n}) > 0$ and $\rho_{\lambda_0}(f_\theta, h_{\theta,n}) \to 0$ $(n \to \infty)$. From what we have just proved, it is easy to see that such $h_{\theta,n}$ exist. Define $h^*_{\theta,n} = E_{\lambda_0}(h_{\theta,n})^{-1} h_{\theta,n}$, $dQ_{\theta,n} = h^*_{\theta,n} d\lambda_0$ and $P_{\theta,n} = Q_{\theta,n}/\mathcal{A}_n$ is clearly a probability measure on $\mathcal{A}_n$. Noting $E_{\lambda_0}(h_{\theta,n}) \to E_{\lambda_0}(f_\theta) = 1$ $(n \to \infty)$, we obtain

$$(10)\qquad \begin{aligned} \|P_\theta - P_{\theta,n}\|_{\mathcal{A}_n} &\leqq \|P_\theta - Q_{\theta,n}\|_{\mathcal{A}} \leqq \rho_{\lambda_0}(f_\theta, h^*_{\theta,n}) \\ &\leqq \rho_{\lambda_0}(f_\theta, h_{\theta,n}) + \rho_{\lambda_0}(h_{\theta,n}, E_{\lambda_0}(h_{\theta,n})^{-1} h_{\theta,n}) \\ &= \rho_{\lambda_0}(f_\theta, h_{\theta,n}) + |1 - E_{\lambda_0}(h_{\theta,n})^{-1}| E_{\lambda_0}(h_{\theta,n}) \\ &\to 0 \qquad (n \to \infty) \end{aligned}$$

for every $\theta \in \Omega$.

The $\mathcal{B}_n$-measurability of $h^*_{\theta,n}$ implies sufficiency of $\mathcal{B}_n$ for $\{P_{\theta,n} | \theta \in \Omega\}$. This, together with (10), implies that $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}$.

**Corollary 1.** *Suppose that $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}$.* If $\rho_{\lambda_0}(E_{\lambda_0}(f_\theta | \mathcal{B}_n), E_{\lambda_0}(f_\theta | \mathcal{B})) \to 0$ $(n \to \infty)$ for every $\theta \in \Omega$, $\mathcal{B}$ is sufficient for $\mathcal{P}$.

Proof. By Theorem 1, we have $\rho_{\lambda_0}(E_{\lambda_0}(f_\theta | \mathcal{B}_n), f_\theta) \to 0$ $(n \to \infty)$ and therefore $f_\theta = E_{\lambda_0}(f_\theta | \mathcal{B}) [\lambda_0]$. This shows that $\mathcal{B}$ is sufficient for $\mathcal{P}$.

**Corollary 2.** *Suppose that $\{\mathcal{B}_n\}$ is approximately sufficient. Then there exist probability measures $P_{\theta,n}$ on $\mathcal{A}$ $(\theta \in \Omega, n = 1, 2, \cdots)$ having the following properties.*

(i) *$\mathcal{B}_n$ is sufficient for $\{P_{\theta,n} | \theta \in \Omega\}$.*

(ii) $\|P_\theta - P_{\theta,n}\|_{\mathcal{A}} \to 0$ $(n \to \infty)$

(iii) $\|P_\theta - P_{\theta,n}\|_{\mathcal{B}_n} = 0$ $(n = 1, 2, \cdots)$.

Proof. Define $dP_{\theta,n} = E_{\lambda_0}(f_\theta | \mathcal{B}_n) d\lambda_0$. Since $\rho_{\lambda_0}(f_\theta, E_{\lambda_0}(f_\theta | \mathcal{B}_n)) \to 0$ $(n \to \infty)$ by Theorem 1, we have $\|P_\theta - P_{\theta,n}\|_{\mathcal{A}} \to 0$ $(n \to \infty)$. (i) and (iii) are clear from the definition of $P_{\theta,n}$.

**Corollary 3.** *Suppose that $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}$. If $\lambda_0$-liminf $\mathcal{B}_n \subset \lambda_0$-liminf $C_n$, $\{C_n\}$ is also approximately sufficient for $\mathcal{P}$.*

Proof. This corollary is clear from (a)$\Leftrightarrow$(d) in Theorem 1 and we omit the proof.

REMARK 1. In [5] $\lambda_0$-liminf $\mathcal{B}_n$ is characterized as the $\sigma$-algebra $\mathcal{B}_0$ having the following properties.

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

$$\text{(A)} \qquad \liminf_{n\to\infty} \int |E_{\lambda_0}(f\,|\,\mathcal{B}_n)|\,d\lambda_0 \geqq \int |E_{\lambda_0}(f\,|\,\mathcal{B}_0)|\,\lambda_0\,d\lambda_0$$

for every bounded $\mathcal{A}$-measurable $f$, and

(ii) any $\sigma$-algebra $\mathcal{B}$ satisfying (A) is contained in $\mathcal{B}_0$. $\lambda_0$-limsup $\mathcal{B}_n$ is also defined there. A $\sigma$-algebra $\tilde{\mathcal{B}}$ is denoted by $\lambda_0$-limsup $\mathcal{B}_n$ if

(i)' $\tilde{\mathcal{B}}$ satisfies

$$\text{(B)} \qquad \limsup_{n\to\infty} \int |E_{\lambda_0}(f\,|\,\mathcal{B}_n)|\,d\lambda_0 \leqq \int |E_{\lambda_0}(f\,|\,\tilde{\mathcal{B}})|\,d\lambda_0$$

for every bounded $\mathcal{A}$-measurable $f$, and

(ii)' any $\sigma$-algebra $\mathcal{B}$ satisfying (B) contains $\tilde{\mathcal{B}}$.

It is proved that, if $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}$, $\lambda_0$-limsup $\mathcal{B}_n$ is sufficient for $\mathcal{P}$ ([4] Theorem 1). Since $\lambda_0$-liminf $\mathcal{B}_n \subset \lambda_0$-limsup $\mathcal{B}_n$ ([5] Theorem 3.4), our result (a)$\Leftrightarrow$(d) in Theorem 1 is an improvement though the assumption (A3) is necessary.

REMARK 2. From Theorem 1 the following question will naturally arise. If there exists $\{P_{\theta,n}\,|\,\theta\in\Omega\}$ on $\mathcal{A}_n$ $(n=1, 2, \cdots)$ such that $\|P_\theta - P_{\theta,n}\|_{\mathcal{A}_n} \to 0$ $(n\to\infty)$ for every $\theta$ and that $\mathcal{B}_n$ is minimal sufficient for $\{P_{\theta,n}\,|\,\theta\in\Omega\}$, is $\lambda_0$-liminf $\mathcal{B}_n$ minimal sufficient? The answer to this question is negative as shown by a very simple counterexample: $X=[0, 1]$, $\mathcal{A}$: Borel field on $[0, 1]$, $\nu$: Lebesgue measure on $\mathcal{A}$, $\mathcal{B}_n=\mathcal{A}_n=\mathcal{A}$ $(n=1, 2, \cdots)$, $P_1=P_2=\nu$. We define $f_{1,n}(x)=\frac{1}{n}x+1-\frac{1}{2n}$, $f_{2,n}(x)=-\frac{1}{n}x+1+\frac{1}{2n}$. Clearly we have $\|f_{1,n}d\nu-P_1\|_{\mathcal{A}_n} \to 0$, $\|f_{2,n}d\nu-P_2\|_{\mathcal{A}_n}\to 0$ and $\nu=\frac{1}{2}f_{1,n}d\nu+\frac{1}{2}f_{2,n}d\nu$. It is easy to see that the smallest $\sigma$-algebra with respect to which $f_{1,n}$, $f_{2,n}$ are measurable is $\mathcal{A}$ itself. Hence $\mathcal{B}_n$ $(=\mathcal{A})$ is minimal sufficient for $\{f_{1,n}d\nu, f_{2,n}d\nu\}$. But $\hat{\mathcal{B}}=\{X, \phi\}$ is sufficient for $\{P_1, P_2\}$. So $\nu$-liminf $\mathcal{B}_n=\mathcal{A}$ is not minimal sufficient.

## 3. Pairwise approximate sufficiency

In this section we shall give an alternative characterization of approximate sufficiency by pairwise approximate sufficiency.

**Theorem 2.** *Under the same condition as in Theorem* 1, if $\{\mathcal{B}_n\}$ *is approxi-*

*mately sufficient for any pair of two* $P_1$, $P_2$ *in* $\mathcal{P}$, *then* $\{\mathcal{B}_n\}$ *is approximately sufficient for* $\mathcal{P}$.

Proof. We divide the proof into the several steps.

The first step. We shall show that it suffices to prove approximate sufficiency of $\{\mathcal{B}_n\}$ for $\mathcal{P}^*$, the dense subset of $\mathcal{P}$. If $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}^*$, we have $\tilde{\rho}_{\lambda_0}\left(\frac{dP}{d\lambda_0}, L_{\lambda_0}(\mathcal{B}_n)\right) \to 0$ $(n\to\infty)$ for every $P\in\mathcal{P}^*$. As we have stated in the proof of $(a)\Rightarrow(b)$ in Theorem 1, we have $\tilde{\rho}_{\lambda_0}\left(\frac{dP}{d\lambda_0}, L_{\lambda_0}(\mathcal{B}_n)\right) \to 0$ $(n\to\infty)$ for every $P\in\mathcal{P}$. Hence, by $(b)\Rightarrow(a)$ in Theorem 1, $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}$.

The second step. We shall prove that $\{\mathcal{B}_n\}$ is approximately sufficient for any finite subset $\{P_1, P_2, \cdots, P_m\}$ of $\mathcal{P}^*$. For this purpose we use the mathematical induction with respect to $m$. Under the assumption that, for $l\leqq k$, $\{\mathcal{B}_n\}$ is approximately sufficient for any $\{P_1, P_2, \cdots, P_l\}$ in $\mathcal{P}^*$, we prove that $\{\mathcal{B}_n\}$ is approximately sufficient for any $\{P_1, P_2, \cdots, P_{k+1}\}$ in $\mathcal{P}^*$. Let $\mu=\sum_{i=1}^{k+1} P_i$. By $(a)\Leftrightarrow(b)$ in Theorem 1, it suffices to show $\tilde{\rho}_\mu\left(\frac{dP_i}{d\mu}, L_\mu(\mathcal{B}_n)\right)\to 0$ $(n\to\infty)$ for every $i=1, 2, \cdots, k+1$, and in particular to show $\tilde{\rho}_\mu\left(\frac{dP_1}{d\mu}, L_\mu(\mathcal{B}_n)\right)\to 0$ $(n\to\infty)$ since the proof of the case $i\neq 1$ is quite analogous. Put $\mu_1=\sum_{i=1}^{k} P_i$, $\mu_2=P_1+P_{k+1}$ and $f_1=\frac{dP_1}{d\mu_1}$, $f_2=\frac{dP_1}{d\mu_2}$. By assumption we have $\tilde{\rho}_{\mu_1}(f_1, L_{\mu_1}(\mathcal{B}_n))\to 0$ $(n\to\infty)$ and $\tilde{\rho}_{\mu_2}(f_2, L_{\mu_2}(\mathcal{B}_n))\to 0$ $(n\to\infty)$. So there exist $\{g_n\}$ and $\{h_n\}$ such that $g_n\in L_{\mu_1}(\mathcal{B}_n)$, $h_n\in L_{\mu_2}(\mathcal{B}_n)$ and $\rho_{\mu_1}(f_1, g_n)\to 0$, $\rho_{\mu_2}(f_2, h_n)\to 0$. Since $0\leqq f_1, f_2\leqq 1$, we can take $g_n$, $h_n$ such that $0\leqq g_n$, $h_n\leqq 1$. Define $\bar{g}_n=\max\left\{g_n, \frac{1}{n}\right\}$, $\bar{h}_n=\max\left\{h_n, \frac{1}{n}\right\}$. It is clear that $\rho_{\mu_1}(f_1, \bar{g}_n)\to 0$ and $\rho_{\mu_2}(f_2, \bar{h}_n)\to 0$. Hence there exists a monotone increasing sequence $\{n_i\}$ of positive integers such that $\bar{g}_{n_i}\to f_1$ (*a.e.* $\mu_1$) and $\bar{h}_{n_i}\to f_2$ (*a.e.* $\mu_2$).

We have

$$\begin{aligned} \frac{dP_1}{d\mu} &= \frac{f_1 f_2}{f_1+f_2-f_1 f_2} && \text{if } f_1 f_2>0 \\ &= 0 && \text{if } f_1=0 \text{ and } f_2=0 \end{aligned}$$

([12] p. 136). Without loss of generality we determine $f_1, f_2$ such that $\{f_1>0, f_2=0\}=\{f_1=0, f_2>0\}=\phi$. Put $\psi_n=\frac{\bar{g}_n\bar{h}_n}{\bar{g}_n+\bar{h}_n-\bar{g}_n\bar{h}_n}$. $\psi_n$ is well-defined because $0<\bar{g}_n$, $\bar{h}_n\leqq 1$. Noting $\mu\approx\mu_1$, $\mu\approx\mu_2$ on $\{f_1 f_2>0\}$, we have

$$\psi_{n_i}\to\frac{dP_1}{d\mu} \quad a.e.\ \mu \quad \text{on} \quad \{f_1 f_2>0\}\ . \tag{11}$$

For $x \in \{\bar{g}_{n_i}\bar{h}_{n_i} \to 0\}$, it is easy to see $\psi_{n_i}(x) \to 0$ $(n \to \infty)$. We have therefore

$$\psi_{n_i}(x) \to \frac{dP_1}{d\mu}(x) \tag{12}$$

for all $x \in \{\bar{g}_{n_i}\bar{h}_{n_i} \to 0\} \cap \{f_1=0 \text{ and } f_2=0\}$.

Since $\mu_i[\{\bar{g}_{n_i}\bar{h}_{n_i} \nrightarrow 0\} \cap \{f_1=0 \text{ and } f_2=0\}]=0$ $(i=1, 2)$,

$$\mu[\{\bar{g}_{n_i}\bar{h}_{n_i} \nrightarrow 0\} \cap \{f_1=0 \text{ and } f_2=0\}]=0\,. \tag{13}$$

It follows from (11)~(13) that $\psi_{n_i} \to \frac{dP_1}{d\mu}$ (*a.e.* $\mu$). Since $|\psi_{n_i} - \frac{dP_1}{d\mu}| \leqq 1$, by Lebesgue's bounded convergence theorem we have $\rho_\mu\left(\frac{dP_1}{d\mu}, \psi_{n_i}\right) \to 0$ $(i \to \infty)$. Since $\psi_{n_i}$ is $\mathcal{B}_{n_i}$-measurable and bounded, we have $\psi_{n_i} \in L_\mu(\mathcal{B}_{n_i})$. So $\tilde{\rho}_\mu\left(\frac{dP_1}{d\mu}, L_\mu(\mathcal{B}_{n_i})\right) \to 0$. By quite a similar to given above, we can prove that, for any subsequence $\{m_n\}$ of $\{n\}$, there exists $\{l_i\} \subset \{m_i\}$ such that $\tilde{\rho}_\mu\left(\frac{dP_1}{d\mu}, L_\mu(\mathcal{B}_{l_i})\right) \to 0$ $(i \to \infty)$. This shows $\tilde{\rho}_\mu\left(\frac{dP_1}{d\mu}, L_\mu(\mathcal{B}_n)\right) \to 0$ $(n \to \infty)$. Thus $\{\mathcal{B}_n\}$ has been shown to be approximately sufficient for any finite subset of $\mathcal{P}^*$.

The third step. As the final step we shall prove that $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}^* = \{P_1, P_2, \cdots\}$. Put $\lambda_m = \sum_{i=1}^{m} \beta_i P_i$, $\lambda_0 = \sum_{i=1}^{\infty} \beta_i P_i$ $(\beta_i > 0,\ \sum_{i=1}^{\infty} \beta_i < \infty)$. $||\lambda_m - \lambda_0||_{\mathcal{A}} \to 0$ (as $m \to \infty$) is clear. $\frac{dP_i}{d\lambda_n}$ exists for $n \geqq i$ and $\frac{dP_i}{d\lambda_n} \to \frac{dP_i}{d\lambda_0}$ $(n \to \infty)$ (*a.e.* $\lambda_0$) for every fixed $i$ ([2] p. 136). From this and $\frac{dP_i}{d\lambda_n} \leqq \beta_i^{-1}$ $(n=0, i, i+1, \cdots)$, we get

$$\rho_{\lambda_0}\left(\frac{dP_i}{d\lambda_n}, \frac{dP_i}{d\lambda_0}\right) \to 0 \qquad (n \to \infty) \tag{14}$$

for every $i$. Since $\{\mathcal{B}_n\}$ is approximately sufficient for $\{P_1, \cdots, P_n\}$ we have $\tilde{\rho}_{\lambda_n}\left(\frac{dP_i}{d\lambda_n}, L_{\lambda_n}(\mathcal{B}_k)\right) \to 0$ $(k \to \infty)$ for every $i$, $n$ with $n \geqq i$. Hence there exists $\{h_{k,n,i}\}$ such that

$$h_{k,n,i} \in L_{\lambda_n}(\mathcal{B}_k),\ \rho_{\lambda_n}\left(\frac{dP_i}{d\lambda_n}, h_{k,n,i}\right) \to 0 \qquad (k \to \infty)\,. \tag{15}$$

Since $\frac{dP_i}{d\lambda_n} \leqq \beta_i^{-1}$, we can assume $0 \leqq h_{k,n,i} \leqq \beta_i^{-1}$ and hence $h_{k,n,i} \in L_{\lambda_0}(\mathcal{B}_k)$. Let $\varepsilon$ be a positive number. We choose $n_0$ such that $||\lambda_{n_0} - \lambda_0||_{\mathcal{A}} < \varepsilon$ and $\rho_{\lambda_0}\left(\frac{dP_i}{d\lambda_0}, \frac{dP_i}{d\lambda_{n_0}}\right) < \varepsilon$. It follows from (15) that there exists $k_0$ such that

$$\rho_{\lambda_{n_0}}\left(\frac{dP_i}{d\lambda_{n_0}}, h_{k,n_0,i}\right)<\varepsilon \text{ for } k\geqq k_0.$$

(16)
$$\left|\rho_{\lambda_{n_0}}\left(\frac{dP_i}{d\lambda_{n_0}}, h_{k,n_0,i}\right)-\rho_{\lambda_0}\left(\frac{dP_i}{d\lambda_{n_0}}, h_{k,n_0,i}\right)\right|$$
$$=\left|\int\left|\frac{dP_i}{d\lambda_{n_0}}-h_{k,n_0,i}\right| d\lambda_{n_0}-\int\left|\frac{dP_i}{d\lambda_{n_0}}-h_{k,n_0,i}\right| d\lambda_0\right|$$
$$\leqq 2\beta_i^{-1}\|\lambda_{n_0}-\lambda_0\|_{\mathcal{A}}<2\beta_i^{-1}\varepsilon .$$

Hence we have for $k\geqq k_0$

$$\rho_{\lambda_0}\left(\frac{dP_i}{d\lambda_0}, h_{k,\,^{n}{}_0,i}\right)\leqq\rho_{\lambda_0}\left(\frac{dP_i}{d\lambda_0}, \frac{dP_i}{d\lambda_{n_0}}\right)+\rho_{\lambda_0}\left(\frac{dP_i}{d\lambda_{n_0}}, h_{k,n_0,i}\right)$$
$$<\rho_{\lambda_0}\left(\frac{dP_i}{d\lambda_0}, \frac{dP_i}{d\lambda_{n_0}}\right)+\rho_{\lambda_{n_0}}\left(\frac{dP_i}{d\lambda_{n_0}}, h_{k,n_0,i}\right)+2\beta_i^{-1}\varepsilon$$
$$<\varepsilon+\varepsilon+2\beta_i^{-1}\varepsilon .$$

Consequently we have $\rho_{\lambda_0}\left(\frac{dP_i}{d\lambda_0}, h_{k,n_0,i}\right)\to 0$ $(k\to\infty)$ for every fixed $i$, which shows $\tilde{\rho}_{\lambda_0}\left(\frac{dP_i}{d\lambda_0}, L_{\lambda_0}(\mathcal{B}_k)\right)\to 0$ $(k\to\infty)$ for every $i$. By $(a)\Leftrightarrow(d)$ in Theorem 1 we see that $\{\mathcal{B}_n\}$ is approximately sufficient for $\mathcal{P}^*$. Thus the proof has been completed.

WASEDA UNIVERSITY

---

## References


[1] A. Berger: *Remark on separable space of probability measures*, Ann. Math. Statist. **22** (1951), 119-120.

[2] P.R. Halmos: Measure Theory, D. van Nostrand Co., New York, 1950.

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

[4] H. Kudō: *On an approximation to a sufficient statistic including a concept of asymptotic sufficiency*, J. Fac. Sci. Univ. Tokyo, Sec. IA **17** (1970), 273–290.

[5] H. Kudō: *A note on the strong convergence of σ-algebras*, Ann. Probability **2** (1974), 76–83.

[6] H. Kudō: *Correction to "On an approximation to a sufficient statistic including a concept of asymptotic sufficiency"*, J. Fac. Sci. Univ. Tokyo, Sec. IA **22** (1975), 449.

[7] J. Neveu: Mathematical Foundations of the Calculus of Probability, Holden-Day, San Francisco, 1965.