TOKYO J. MATH.
VOL. 13, NO. 1, 1990

Remarks on Bayes Sufficiency

Sakutarō YAMADA and Takeru SUZUKI

Tokyo University of Fisheries and Waseda University

Introduction.

It is well known that there are various definitions of "sufficiency". In addition to the usual definition of sufficiency represented by the existence of a common conditional probability, we have the notions of pairwise sufficiency, PSS (pairwise sufficiency with supports), test sufficiency and Bayes sufficiency. These notions coincide with one another in the dominated case. Ramamoorthi ([6]), and Roy and Ramamoorthi ([7]), discussed Bayes sufficiency in undominated cases, but most results are restricted to the countably generated subfield cases.

In this note we show, by examples, that PSS does not imply Bayes sufficiency in case of a continuous a priori distribution (cf. [5]), and the existence of the smallest PSS which is Bayes sufficient. This latter example shows the result by Kusama and Fujii ([4]) does not hold if we replace test sufficiency by Bayes sufficiency.

By a statistical experiment we mean a triplet $(\mathscr{X}, \mathscr{A}, \mathscr{P})$, where $\mathscr{P}=\{P_\theta; \theta\in\Theta\}$ is a family of probability measures on $(\mathscr{X}, \mathscr{A})$. Θ is referred to as the parameter space of the experiment. Let $\mathscr{C}$ be a sigma-field of subsets of Θ which includes the sigma-field generated by the family of mappings defined by $\theta\in\Theta\to P_\theta(A)$, $A\in\mathscr{A}$. For any a priori distribution λ on $(\Theta, \mathscr{C})$, let $\mathscr{P}*\lambda$ be a probability measure on $(\mathscr{X}\times\Theta, \mathscr{A}\times\mathscr{C})$ defined by

$$\mathscr{P}*\lambda(A\times C)=\int_C P_\theta(A)d\lambda(\theta) .$$

Let $\mathscr{B}$ be a sub-sigma-field of $\mathscr{A}$ and Λ be a family of a priori distributions on $(\Theta, \mathscr{C})$.

DEFINITION. $\mathscr{B}$ is called Bayes sufficient w.r.t. Λ if, for any $\lambda\in\Lambda$, $\mathscr{A}\times\Theta$ and $\mathscr{X}\times\mathscr{C}$ are conditionally independent given $\mathscr{B}\times\Theta$ w.r.t.

Received June 22, 1989
Revised October 1, 1989

$\mathscr{P}*\lambda$. If $\mathscr{B}$ is Bayes sufficient w. r. t. the family of all a priori distributions then we simply say that $\mathscr{B}$ is Bayes sufficient. Here $\mathscr{A}\times\Theta$ denotes the sigma-field of sets of the form $A\times\Theta$, $A\in\mathscr{A}$. $\mathscr{B}\times\Theta$, $\mathscr{X}\times\mathscr{C}$ are defined similarly.

1. Results.

This section is separated into four parts. First let's consider the relationship between Bayes sufficiency and pairwise sufficiency. The next proposition is a slight modification of one part of Proposition 1.7 of [5]. This Proposition 1.7 also shows that Bayes sufficiency implies pairwise sufficiency.

PROPOSITION 1. *A sub-sigma-field $\mathscr{B}$ is pairwise sufficient if and only if $\mathscr{B}$ is Bayes sufficient w. r. t. the totality of discrete a priori distributions on $(\Theta, \mathscr{C})$.*

PROOF. One part of Proposition 1.7 of [5] shows that pairwise sufficiency is equivalent to Bayes sufficiency w. r. t. the totality of convex combinations of two point distributions. Hence the only if part is to be proved. Let λ be any discrete a priori distribution on $\{\theta_1, \theta_2, \cdots\}$ with probabilities $\{a_1, a_2, \cdots\}$ and let $\mathscr{P}_\lambda=\{P_{\theta_i}; i\geqq 1\}$. Then $\mathscr{B}$ is sufficient for $(\mathscr{X}, \mathscr{A}, \mathscr{P}_\lambda)$. Hence $\mathscr{B}$ is Bayes sufficient for the parameter space $(\{\theta_1, \theta_2, \cdots\}, 2^{\{\theta_1,\theta_2,\cdots\}})$. Let $\hat{\lambda}$ be the restriction of λ on $\{\theta_1, \theta_2, \cdots\}$. Then $\mathscr{A}\times\{\theta_1, \theta_2, \cdots\}$ and $\mathscr{X}\times 2^{\{\theta_1,\theta_2,\cdots\}}$ are conditionally independent given $\mathscr{B}\times\{\theta_1, \theta_2, \cdots\}$ w. r. t. $\mathscr{P}*\hat{\lambda}$. Hence for any bounded $\mathscr{A}$-measurable function f there exists a $\mathscr{B}$-measurable function g such that

$$\int_{B\times C}\bar{f}(x, \theta)d\mathscr{P}*\hat{\lambda}=\int_{B\times C}\bar{g}(x, \theta)d\mathscr{P}*\hat{\lambda} \qquad \text{for all } B\in\mathscr{B} \text{ and } C\subset\{\theta_1, \theta_2, \cdots\},$$

where $\bar{f}(x, \theta)=f(x)$, $\bar{g}(x, \theta)=g(x)$. The left-hand side is equal to

$$\int_C\int_B f(x)dP_{\theta_i}d\hat{\lambda}=\sum\left\{a_i\int_B f(x)dP_{\theta_i};\ \theta_i\in C\right\}.$$

The right-hand side is equal to

$$\int_C\int_B g(x)dP_{\theta_i}d\hat{\lambda}=\sum\left\{a_i\int_B g(x)dP_{\theta_i};\ \theta_i\in C\right\}.$$

Take any $C\in\mathscr{C}$. Then we have

$$\int_{B\times C}\bar{f}(x, \theta)d\mathscr{P}*\lambda=\int_C\int_B f(x)dP_\theta d\lambda$$

$$=\sum\left\{a_i\int_B f(x)dP_{\theta_i};\ \theta_i\in C\right\}=\sum\left\{a_i\int_B g(x)dP_{\theta_i};\ \theta_i\in C\right\}$$
$$=\int_C\int_B g(x)dP_\theta d\lambda=\int_{B\times C}\bar{g}(x,\ \theta)d\mathscr{P}*\lambda\ .$$

This shows $\mathscr{A}\times\Theta$ and $\mathscr{X}\times\mathscr{C}$ are conditionally independent given $\mathscr{B}\times\Theta$ w. r. t. $\mathscr{P}*\lambda$. Hence B is Bayes sufficient w. r. t. the totality of discrete a prioi distributions.

The next example shows that pairwise sufficiency does not imply Bayes sufficiency w. r. t. a single continuous a priori distribution, and that PSS does not imply Bayes sufficiency.

EXAMPLE 1. Let $\mathscr{X}=\boldsymbol{R}^1$, $\mathscr{A}=$Borel field, $\Theta=\boldsymbol{R}^1-\{0\}$, $\mathscr{P}=\{P_\theta;\ \theta\in\Theta\}$ and P_θ be the point probability measure at θ. Let $\mathscr{C}$ be the Borel field of subsets of Θ. This $\mathscr{C}$ is equal to the sigma-algebra generated by all mappings defined by $\theta\to P_\theta(A)$, $A\in\mathscr{A}$. Let $\mathscr{B}$ be the sub-sigma-field of $\mathscr{A}$ generated by all single points but 0. Then $\mathscr{B}$ is the smallest PSS ([8] Proposition 1). Suppose that $\mathscr{B}$ is Bayes sufficient. Then for any a priori distribution λ on $(\Theta,\ \mathscr{C})$ and $C\in\mathscr{C}$, there exists a $\mathscr{B}$-measurable function g such that

$$\int_{A\in\Theta}\bar{g}(x,\ \theta)d\mathscr{P}*\lambda=\mathscr{P}*\lambda((A\times\Theta)\cap(\mathscr{X}\times\mathscr{C}))\qquad\text{for all}\quad A\in\mathscr{A}\ .$$

Then the left-hand side is equal to

$$\int_\Theta\int_A g(x)dP_\theta(x)d\lambda(\theta)=\int_\Theta\int_{A-\{0\}}g(x)dP_\theta(x)d\lambda(\theta)$$
$$=\int_\Theta g(\theta)I_{A-\{0\}}(\theta)d\lambda(\theta)=\int_{A-\{0\}}g(\theta)d\lambda(\theta)\ ,$$

and the right-hand side is equal to

$$\mathscr{P}*\lambda(A\times C)=\int_C P_\theta(A)d\lambda(\theta)=\int_C P_\theta(A-\{0\})d\lambda(\theta)$$
$$=\int_C I_{A-\{0\}}(\theta)d\lambda(\theta)=\int_{A-\{0\}}I_C(\theta)d\lambda(\theta)\ .$$

Hence we have

$$g(\theta)=I_C(\theta)\qquad\text{a.e. }\lambda\ . \tag{1}$$

Let's take as λ a continuous a priori distribution satisfying

$$\lambda([-2,\ -1])>0\ ,\quad \lambda([1,\ 2])>0\ ,\quad \lambda(\Theta-C)>0\qquad\text{for}\quad C=[-2,\ -1]\cup[1,\ 2]\ .$$

Then by (1),

$$\Theta \cap [g \neq I_C] = \{[g \neq 1] \cap C\} \cup \{[g \neq 0] \cap (\Theta - C)\} \qquad \text{a.e. } \lambda .$$

Case 1. $g(0)=0$. As $0 \in [g \neq 1]$ and g is B-measurable, $[g=1]$ is countable. So $\lambda([g \neq 1] \cap C) = \lambda(C) > 0$. This implies $\lambda(\Theta \cap [g = I_C]) > 0$ and contradicts (1).

Case 2. $g(0)=1$. As $0 \in [g \neq 0]$ and g is $\mathscr{B}$-measurable, $[g=0]$ is countable. So $\lambda([g \neq 0] \cap (\Theta - C)) = \lambda(\Theta - C) > 0$. This implies $\lambda(\Theta \cap [g \neq I_C]) > 0$ and contradicts (1).

Case 3. $g(0) \neq 0, 1$. As $0 \in [g \neq 1] \cap [g \neq 0]$ and g is $\mathscr{B}$-measurable, $[g=1] \cup [g=0]$ is countable. So $\lambda([g \neq 1] \cap [g=0] \cap C) = \lambda(C) > 0$. This implies $\lambda(\Theta \cap [g \neq I_C]) > 0$ and contradicts (1).

Hence $\mathscr{B}$ is not Bayes sufficient. It is easily proved that $\mathscr{B}$ is not test sufficient.

REMARK 1. If we take, in Example 1, $\mathscr{A} = 2^{\mathscr{X}}$, then $\mathscr{C}$ must be equal to 2^{Θ}. Hence, under the continuum hypothesis, any a priori distribution on $(\Theta, \mathscr{C})$ is discrete, and hence $\mathscr{B}$ is Bayes sufficient because it is pairwise sufficient for $(\mathscr{X}, 2^{\mathscr{X}}, \mathscr{P})$.

The second part of this section is related to the result of Kusama and Fujii ([4]). It shows that the smallest PSS is not test sufficient when the underlying experiment is weakly dominated and not dominated. The next example shows that in the above result we can't replace test sufficiency with Bayes sufficiency.

EXAMPLE 2. Let $\mathscr{X}$ be any set with non-measurable cardinal and $\mathscr{A} = 2^{\mathscr{X}}$. For any $\theta \in \Theta = \mathscr{X}$, P_θ denotes the point probability measure at θ. Let $\mathscr{P} = \{P_\theta;\ \theta \in \Theta\}$ and $\mathscr{B}$ be the sub-sigma-field of $\mathscr{A}$ generated by all single points of $\mathscr{X}$. To discuss Bayes sufficiency it is noted that $\mathscr{C} = 2^{\Theta}$. In this case, because $\mathscr{X}$ is non-measurable, any probability measure on $(\Theta, \mathscr{C})$ is discrete. Hence $\mathscr{B}$ is Bayes sufficient by Proposition 1 and because it is pairwise sufficient. To put it more precisely, $\mathscr{B}$ is the smallest PSS ([3], Example 4.1). $\mathscr{B}$ is not test sufficient ([4]).

REMARK 2. Brown ([2]) showed, in the discrete case, any test sufficient sub-sigma-field includes a sufficient sub-sigma-field. Hence in this case test sufficiency implies Bayes sufficiency. Our Example 2 shows the converse does not hold. In general does test sufficiency imply Bayes sufficiency? This is a weaker version of the general Brown's problem and is still an open problem. Ramamoorthi ([6], Theorem 2.2.1) showed that, if all the sigma-fields concerned are countably generated, test sufficiency implies Bayes sufficiency.

The third part of this section is related to Bayes sufficiency for completed sigma-fields. This gives an analogue to Proposition 1. For any sub-sigma-field $\mathscr{B}$ of $\mathscr{A}$ we define its pairwise completion by $\tilde{\mathscr{B}}(\mathscr{A})=\{A\in\mathscr{A};$ For any $P, Q\in\mathscr{P}$ there exists $B_{P,Q}\in\mathscr{B}$ satisfying $P(A\triangle B_{P,Q})=Q(A\triangle B_{P,Q})=0\}$.

Proposition 2. *Let $(\mathscr{X}, \mathscr{A}, \mathscr{P})$ be weakly dominated. Then $\mathscr{B}$ is pairwise sufficient if and only if $\tilde{\mathscr{B}}(\mathscr{A})$ is Bayes sufficient.*

This follows easily from Corollary 2 of [9] and we omit the proof. The next example shows in Proposition 2 we can't replace the pairwise completion with the weak completion of $\mathscr{B}$ which is defined by $\hat{\mathscr{B}}(\mathscr{A})=\{A\in\mathscr{A};$ For any $P\in\mathscr{P}$ there exists a $B_p\in\mathscr{B}$ such that $P(A\triangle B_p)=0\}$.

Example 3. Let $\mathscr{X}=\boldsymbol{R}^1$, $\mathscr{A}=2^{\mathscr{X}}$, $\mathscr{P}$ be the totality of point probability measures. Let $\mathscr{B}$ be the sub-sigma-field of $\mathscr{A}$ generated by all single points but 0 and 1. Then it is easy to prove that $\mathscr{B}$ is not pairwise sufficient and $\hat{\mathscr{B}}(\mathscr{A})=\mathscr{A}$, which implies $\hat{\mathscr{B}}(\mathscr{A})$ is Bayes sufficient. This example is the same to the one in Remark 2 of [9].

Finally we remark that there exist some connections between Bayes sufficiency and the existence of consistent estimates. Let $(\mathscr{X}, \mathscr{A}, \mathscr{P})$, $\mathscr{P}=\{P_\theta; \theta\in\Theta\}$, be a given statistical experiment, $\mathscr{B}$ be a sub-sigma-field of $\mathscr{A}$, and let Λ be a family of a priori distributions on $(\Theta, \mathscr{C})$. It is easily seen from the equivalent formulation of Bayes sufficiency (Ramamoorthi [6], Proposition 2.1.1, or Roy and Ramamoorthi [7], Proposition 2.1) that $\mathscr{B}$ is Bayes sufficient w. r. t. Λ if for every $\lambda\in\Lambda$ and every real valued $\mathscr{C}$-measurable function $h(\theta)$ on Θ there exists $\mathscr{B}$-measurable function $f(x)$ on $\mathscr{X}$ such that

$$P_\theta\{x; f(x)=h(\theta)\}=1\ , \qquad \lambda\text{-a.e.}$$

In other words, the existence of $\mathscr{B}$-measurable (λ-)consistent estimates for every $\lambda\in\Lambda$ and every $\mathscr{C}$-measurable function $h(\theta)$ implies Bayes sufficiency of $\mathscr{B}$ w. r. t. Λ. Hence, in particular, if there exist $\mathscr{B}$-measurable consistent estimates in a usual sense for every $h(\theta)$ then $\mathscr{B}$ is Bayes sufficient. It is not known whether the converse of this statement holds or not. When $\Lambda=\{\theta\}$ is one point set, some interesting relations between the existence of $\mathscr{A}$-measurable (λ-)consistent estimates and weak zero-one sets are discussed in Breiman et al. ([1]). In Roy and Ramamoorthi ([7]) they discuss some relationship between existence of a 'nearly' consistent estimator, which is called a measurable estimator in their paper, and measurable coherency.

ACKNOWLEDGEMENTS. We would like to thank the referee for carefully reading the paper and for many useful comments.

References

[1] L. BREIMAN, L. LECAM and L. SCHWARTZ, Consistent estimates and zero-one sets, Ann. Math. Statist., **35** (1964), 157-161.

[2] L. BROWN, On a theorem of Morimoto concerning sufficiency for discrete distributions, Ann. Math. Statist., **3** (1975), 1180-1182.

[3] J. K. GHOSH, H. MORIMOTO and S. YAMADA, Neyman factorization and minimality of pairwise sufficient subfields, Ann. Statist., **9** (1981), 514-530.

[4] T. KUSAMA and J. FUJII, A note on test sufficiency in weakly dominated statistical experiments, Tokyo J. Math., **10** (1987), 133-136.

[5] F. MARTIN, J.-L. PETIT and M. LITTAYE, Indépendance conditionnelle dans le modèle statistique bayésien, Ann. Inst. H. Poincaré Sect. B, **9** (1973), 19-40.

[6] R. V. RAMAMOORTHI, Sufficiency, pairwise sufficiency and Bayes sufficiency in undominated experiments, Thesis, Indian Statistical Institute, 1980.

[7] K. K. ROY and R. V. RAMAMOORTHI, Relationship between Bayes, classical and decision theoretic sufficiency, Sankhyā Ser. A, **41** (1979), 48-58.

[8] E. SIEBERT, Pairwise sufficiency, Z. Wahrsch. Verw. Gebiete, **46** (1979), 237-246.

[9] S. YAMADA, On completion of sigma-fields and pairwise sufficiency, Sankhyā Ser. A, **42** (1980), 185-200.

Present Address:
SAKUTARŌ YAMADA
DEPARTMENT OF FISHERIES, RESOURCE MANAGEMENT, TOKYO UNIVERSITY OF FISHERIES
KONAN, MINATO-KU, TOKYO 108, JAPAN

TAKERU SUZUKI
DEPARTMENT OF MATHEMATICS, SCHOOL OF SCIENCE AND ENGINEERING,
WASEDA UNIVERSITY
OKUBO, SHINJUKU-KU, TOKYO 169, JAPAN