

ON COMPACTNESS OF THE STATISTICAL STRUCTURE AND SUFFICIENCY

TOKITAKE KUSAMA AND SAKUTARŌ YAMADA

(Received November 6, 1970)

(Revised September 17, 1971)

It is well known that the theory of sufficiency in mathematical statistics is closely related to the properties of the statistical structure (X, S, M) , where X is a set, S is a σ -field of subsets of X and M is a collection of probability measures on S . When M is dominated, that is, each measure in M is absolutely continuous with respect to a common probability measure, we have the several important results about sufficiency (for example [1], [2]).

T. S. Pitcher introduced a more general condition than the domination, which we call compactness, and proved the existence of a minimal sufficient sub- σ -field under this condition, which had been guaranteed if M is dominated ([3]).

In §1, we give a necessary and sufficient condition for compactness in the case of discrete probability measures and a necessary condition in general case. In §2 we prove a few results about the relations between sufficiency and compactness. In the final section we give the negative answers to the open problems posed by Pitcher.

It is a pleasure to express our best thanks to Mr. H. Morimoto and Mr. M. Takahashi for their valuable suggestions.

0. In this section, after stating some notations and definitions, we refer to Pitcher's definition of compactness and state some of his results which will be referred to in this paper.

The statistical structure (X, S, M) is to be kept in mind throughout the paper and all σ -fields entering the discussion is implicitly assumed to be a sub- σ -field of S . For a σ -field T and a finite measure P , we define an outer measure $\Gamma(*|T, P)$ by $\Gamma(A|T, P) = \inf_{A \subset B \in T} P(B)$ for any subset A of X .

Let T_1 and T_2 be two σ -fields. We write $T_1 \doteq T_2$ if for any $A \in T_1$ there exists $B \in T_2$ such that $\Gamma(A \Delta B|S, P) = 0$ for all $P \in M$ and vice versa, where $A \Delta B$ denotes the symmetric difference of A and B . For a σ -field T , put

$\rho_T(x) = \bigcap_{x \in A \in T} A$. T will be called an atomic σ -field if $\rho_T(x) \in T$ for all $x \in X$. $\{\rho_T(x) | x \in X\}$ forms a partition of X and we call it the partition induced by T . A σ -field will be said to be complete if it is closed under the formation of unions of arbitrarily many number of sets in it. A complete field is clearly atomic.

For a σ -field T , we define $\hat{T}(S) = \{A | \text{for each } P (\in M), \text{ there exists } D_P \in T \text{ such that } \Gamma(A \Delta D_P | S, P) = 0\}$. We write $f = g[P]$ if f is equal to g almost everywhere with respect to P .

For each $P (\in M)$, S -measurable f and real number p with $1 \leq p < \infty$ we write $\|f\|_{p,P}$ for the number $(\int_X |f|^p dP)^{1/p}$, and $\|f\|_{\infty,P}$ for the P -essential supremum of $|f|$. For all $\infty \geq p \geq 1$ we define $\|f\|_{p,M} = \sup_{P \in M} \|f\|_{p,P}$. Let $E_p(X, S, M)$ be the set of all f with $\|f\|_{p,M} < \infty$. Then $E_p(X, S, M)$ with $\|f\|_{p,M}$ as norm is a Banach space. Let $B_p(X, S, M)$ be the unit ball in $E_p(X, S, M)$.

Now we proceed to define the compactness of (X, S, M) . For each $P \in M$ and $h \in L_q(P)$ $\left(\frac{1}{p} + \frac{1}{q} = 1\right)$, $l(h, P)$ through the formula $l(h, P)(f) = \int_X f h dP$ is a continuous linear functional on $E_p(X, S, M)$. Let $\mathcal{E}_p(X, S, M)$ be the totality of finite linear combinations of a finite number of $l(h, P)$. We consider the weakest topology in which the elements of $\mathcal{E}_p(X, S, M)$ are continuous and call it $\mathcal{E}_p(X, S, M)$ topology. (X, S, M) is said to be compact if $B_p(X, S, M)$ is compact with respect to the $\mathcal{E}_p(X, S, M)$ topology for some p ($1 < p < \infty$). It is known ([3] Theorem 1.1) that if $B_p(X, S, M)$ is $\mathcal{E}_p(X, S, M)$ compact for some p ($1 < p < \infty$) then it is compact for all such p .

Let $W_p(P)$ be the weakly topologized unit ball in $L_p(P)$ and $\prod_{P \in M} W_p(P)$ the direct product space of $W_p(P)$ with usual Tychonoff topology. Elements of $\prod_{P \in M} W_p(P)$ will be denoted by (f_P) . We designate by i_p the diagonal mapping which maps each $f \in B_p(X, S, M)$ to the element of the product space $\prod_{P \in M} W_p(P)$ whose value at $W_p(P)$ is f . According to [3] (X, S, M) is compact if and only if $i_p(B_p(X, S, M))$ is closed in $\prod_{P \in M} W_p(P)$ for some p ($1 < p < \infty$).

Here we state some results obtained by Pitcher which will be often referred to in this paper.

Theorem A. ([3] Lemma 1. 2.) *The following are equivalent:*

1. $(f_P)_{P \in M}$ is in the closure of $i_p(B_p(X, S, M))$;
2. for every finite set P_1, \dots, P_n from M there is an f in $B_p(X, S, M)$ satisfying $f = f_{P_i}[P_i]$ for $i = 1, \dots, n$;
3. for every countable set (P_i) from M there is an f in $B_p(X, S, M)$ satisfying $f = f_{P_i}[P_i]$ for all i .

Theorem B. ([3] Theorem 2. 1.) *If (X, S, M) is compact, then $S \doteq \hat{S}(S)$.*

Theorem C. ([3] Theorem 2. 2.) *If (X, S, M) is compact and T is a sub- σ -field of S , then $(X, \hat{T}(S), M)$ is compact.*

Theorem D. ([3] Theorem 2. 3.) *If T is a sufficient sub- σ -field for $(X, \hat{S}(S), M)$, then $T \doteq \hat{T}(S)$.*

Theorem E. ([3] Theorem 2. 4.) *If $\hat{T}(S)$ is a sufficient sub- σ -field for $(X, \hat{S}(S), M)$ then $(X, \hat{S}(S), M)$ is compact if and only if $(X, \hat{T}(S), M)$ is compact and $\{b | b \in B_\infty(X, \hat{S}(S), M) \text{ and } E(b | \hat{T}(S), M) = 0\}$ is compact in the $\mathcal{E}_1(X, \hat{S}(S), M)$ topology.*

Theorem F. ([3] Theorem 2. 5.) *If (X, S, M) is compact, then there exists a best sufficient sub- σ -field of S , i.e., a sufficient sub- σ -field T such that $T \subset T_1$ for any other sufficient sub- σ -field T_1 .*

1. Throughout this section we assume that S is atomic and let $\{A_\alpha | \alpha \in \Lambda\}$ be the partition of X induced by S . A finite measure P on S is said to be discrete if there exist a countable number of A_α , say $A_{\alpha_1}, A_{\alpha_2}, \dots$ such that $P(\bigcup_{i=1}^{\infty} A_{\alpha_i}) = P(X)$. Let $C(P)$ be the set of all α satisfying $P(A_\alpha) > 0$ and X_P be the set $\bigcup \{A_\alpha | \alpha \in C(P)\}$. For a subset Λ' of Λ let $X(\Lambda')$ be the set $\bigcup \{A_\alpha | \alpha \in \Lambda'\}$.

Theorem 1.1. *Let M be a family of discrete probability measures. Let us assume that for any A_α there exists $P \in M$ with $P(A_\alpha) > 0$. Then the following three assertions are equivalent:*

- (1) (X, S, M) is compact;
- (2) $S = \hat{S}(S)$;
- (3) S is complete.

Proof. (1) \Rightarrow (2). Since (X, S, M) is compact, we have $S \doteq \hat{S}(S)$ (Theorem B). By the assumption of our theorem, $\Gamma(A | S, P) = 0$ for all $P \in M$ implies $A = \phi$. Hence we have $S = \hat{S}(S)$.

(2) \Rightarrow (3). Suppose that S is not complete. Then there exists $\Lambda' (\subset \Lambda)$ such that $X(\Lambda') \notin S$. We put $D_P = X - (X_P - X(\Lambda'))$ for each $P \in M$. Since $X_P \in S$, it is clear $D_P \in S$. We have $X(\Lambda') \Delta D_P = D_P - X(\Lambda') \subset X - X_P \in S$ and $P(X - X_P) = 0$. This implies $\Gamma(X(\Lambda') \Delta D_P | S, P) = 0$ and hence $X(\Lambda') \in \hat{S}(S)$. Thus we have $S \not\subseteq \hat{S}(S)$.

(3) \Rightarrow (1). We first note that an S -measurable f is a constant on each A_α and its converse is also true since S is complete. Let us denote this constant value by f^α . Moreover we note that $f = g[P]$ implies $f^\alpha = g^\alpha$ for each $\alpha \in C(P)$. Let (f_P) be an element in the closure of $i_P(B_P(X, S, M))$. It follows from Theorem A that, for $P, P' \in M$, there exists a g in $B_P(X, S, M)$ such that $g = f_P[P]$ and $g = f_{P'}[P']$. Then by the above notes we have $f_P^\alpha = f_{P'}^\alpha$ for each

$\alpha \in C(P) \cap C(P')$. Hence for each $\alpha \in \Lambda$ there exists a real c_α such that $f_P^\alpha = c_\alpha$ for all P satisfying $\alpha \in C(P)$. We define a function f such that $f(x) = c_\alpha$ for $x \in A_\alpha$. f is clearly S -measurable because S is complete. We have $f = f_P[P]$ for all $P \in M$. Since $(f_P) \in \prod_{P \in M} W_p(P)$, we have $f \in B_p(X, S, M)$ and $i_p(f) = (f_P)$. Hence $(f_P) \in i_p(B_p(X, S, M))$, which shows closedness of $i_p(B_p(X, S, M))$. Hence (X, S, M) is compact.

It is easy to see that a probability measure P can be decomposed as $P = P^1 + P^2$, where P^1 is a discrete measure and P^2 is a measure satisfying $P^2(A_\alpha) = 0$ for all α .

Theorem 1.2. *If the cardinality of Λ is nonmeasurable, the completeness of S implies the compactness of (X, S, M) .*

Proof. By the assumption of our theorem and completeness of S , we have $P^2(X) = 0$ for all $P \in M$. So all P in M are discrete. Hence it follows from Theorem 1.1 that (X, S, M) is compact.

REMARK 1.1. It is known that the cardinality of the continuum is non-measurable under the continuum hypothesis ([4]).

The following theorem gives a necessary condition for compactness of a statistical structure.

Theorem 1.3. *Suppose that (X, S, M) is compact. Then we have $\Gamma(H|S, P^2) = 0$, where $H = \bigcup_{P \in M} X_P$.*

Proof. Suppose that there exists $\tilde{P} \in M$ such that $\Gamma(H|S, \tilde{P}^2) > 0$. Let f_P be the indicator function of X_P . For every finite set P_1, \dots, P_n from M , if g is the indicator function of $\bigcup_{i=1}^n X_{P_i}$, we have $g = f_{P_i}[P_i]$ ($i=1, 2, \dots, n$). This shows $(f_P) \in i_p(B_p(X, S, M))$ (Theorem A). Hence there exists an h in $B_p(X, S, M)$ such that $h = f_P[P]$ for all $P \in M$. h must satisfy $h(x) = 1$ for $x \in A_\omega$ if $A_\omega \subset H$, and hence $H \subset \{x | h(x) = 1\} \in S$. Since $\Gamma(H|S, \tilde{P}^2) > 0$ by assumption, we have $\tilde{P}^2(\{x | h(x) = 1\}) > 0$. Consequently we have $\tilde{P}^2(\{x | h(x) = 1\} - X_{\tilde{P}}) > 0$ and therefore $\tilde{P}(\{x | h(x) = 1\} - X_{\tilde{P}}) > 0$. It is clear that $f_P(x) = 0$ and $h(x) = 1$ on $\{x | h(x) = 1\} - X_{\tilde{P}}$, that is, that f_P and h take different values on a set of positive measure, which is contradictory to $h = f_{\tilde{P}}[\tilde{P}]$. Hence we have $\Gamma(H|S, P^2) = 0$ for all $P \in M$.

REMARK 1.2. To show that there are assertions about sufficiency true in the dominated case but not necessarily true in the undominated case, several authors gave the pathological examples of the statistical structure ([1], [5], [6]). Using Theorem 1.1 or 1.3 it will be verified easily that these structures are not compact.

2. At first we shall state the definitions of the sufficiency and the pairwise sufficiency.

Let T be a sub- σ -field of S . For a probability measure P and a bounded S -measurable function f , there exists a T -measurable function $E_P(f|T)$ satisfying

$$\int_A f \, dP = \int_A E_P(f|T) \, dP$$

for all $A \in T$. This $E_P(f|T)$ is called the conditional expectation function of f given T and P . A sub- σ -field T of S is sufficient for (X, S, M) if corresponding to each bounded S -measurable function f , there exists a T -measurable function φ_f such that

$$\varphi_f = E_P(f|T)[P]$$

for each P in M . We shall denote this φ_f by $E(f|T, M)$. A sub- σ -field T of S is pairwise sufficient for (X, S, M) if it is sufficient for every pair $\{P_1, P_2\}$ of measures in M .

In the dominated case it is well known that if T_1 is sufficient and $T_1 \subset T_2$ then T_2 is also sufficient and that the concept of sufficiency coincides with that of pairwise sufficiency. The aim of this section is to study these results under the compactness condition.

Theorem 2.1. *Suppose that T_1 is a pairwise sufficient sub- σ -field for (X, S, M) and $T_1 \subset T_2 (\subset S)$. Then T_2 is sufficient for (X, S, M) if (X, T_2, M) is compact.*

Proof. Let T be a sub- σ -field. Let $e_P(E|T)$ be the conditional expectation function of $\chi_E(E \in S)$ given T and a probability measure P . Since T_1 is pairwise sufficient for (X, S, M) , so is T_2 . Therefore for every finite set P_1, \dots, P_n from M there is a T_2 -measurable g such that $g = e_{P_i}(E|T_2)[P_i]$ ($i=1, \dots, n$). Since we can take g such that $0 \leq g(x) \leq 1$ for all x , we may assume $g \in B_p(X, T_2, M)$. Accordingly $(e_P(E|T_2))_{P \in M}$ is contained in the closure of $i_p(B_p(X, T_2, M))$ (Theorem A). Hence $(e_P(E|T_2))_{P \in M}$ is in $i_p(B_p(X, T_2, M))$ since (X, T_2, M) is compact. Thus there is an h in $B_p(X, T_2, M)$ such that $h = e_P(E|T_2)[P]$ for all $P (\in M)$, which shows the sufficiency of T_2 .

Corollary 2.1. *Suppose that (X, S, M) is compact and T_1 is pairwise sufficient for (X, S, M) . If $T_1 \subset T_2$ and $\hat{T}_2(S) \subset S$, then $\hat{T}_2(S)$ is sufficient for (X, S, M) .*

Proof. By Theorem C $(X, \hat{T}_2(S), M)$ is compact. Under the conditions of the above Theorem 2.1, $\hat{T}_2(S)$ is sufficient for (X, S, M) .

Corollary 2.2. *Suppose that (X, S, M) is compact and that T_0 is the*

minimal sufficient sub- σ -field and $T_0 \subset T \subset S$. If (X, T, M) is compact, then $K = \{f \mid f \in B_\infty(X, \hat{S}(S), M), E(f \mid \hat{T}(S), M) = 0\}$ is $\mathcal{E}_1(X, \hat{S}(S), M)$ compact.

Proof. By Theorem 2.1 T is sufficient for (X, S, M) . Since (X, S, M) is compact, we have $S \doteq \hat{S}(S)$ (Theorem B). So T is sufficient for $(X, \hat{S}(S), M)$ and therefore $T \doteq \hat{T}(S)$ (Theorem D). Hence $\hat{T}(S)$ is sufficient for $(X, \hat{S}(S), M)$ and therefore $E(f \mid \hat{T}(S), M)$ can be taken independently of $P (\in M)$. On the other hand it follows from $S \doteq \hat{S}(S)$ that $(X, \hat{S}(S), M)$ is also compact. Hence by Theorem E, K is $\mathcal{E}_1(X, \hat{S}(S), M)$ compact.

Theorem 2.2. *If (X, S, M) is compact and if $T (\subset S)$ is sufficient for (X, S, M) , then (X, T, M) is compact.*

Proof. Since (X, S, M) is compact, we have $S \doteq \hat{S}(S)$ (Theorem A). Thus it is shown that T is also sufficient for $(X, \hat{S}(S), M)$. Hence $T \doteq \hat{T}(S)$ holds (Theorem D). $(X, \hat{T}(S), M)$ is compact (Theorem C) and therefore (X, T, M) is compact.

Corollary 2.3. *Suppose that (X, S, M) is compact, that $T_1 \subset T_2 \subset S$ and that T_1 is pairwise sufficient for (X, S, M) . Then the sufficiency of T_2 is equivalent to the compactness of (X, T_2, M) .*

Proof. This is a direct consequence of Theorems 2.1 and 2.2. So we shall omit the proof.

REMARK 2.1. Suppose that (X, S, M) satisfies the conditions of Theorem 1.1 and that S is complete. Under these conditions Basu and Ghosh proved that there exists a minimal sufficient sub- σ -field and Morimoto proved that a pairwise sufficient sub- σ -field is sufficient if and only if S is complete ([7], [8]). (X, S, M) is clearly compact by Theorem 1.1. So the result due to Basu and Ghosh is immediate from the theorem obtained by Pitcher which guarantees the existence of the minimal sufficient sub- σ -field under the compactness condition (Theorem F). The result obtained by Morimoto is easily deduced from Theorem 1.1 and Corollary 2.3.

3. Pitcher mentioned two unsolved problems in his paper.

- (1) If (X, S, M_i) ($i=1, 2$) are compact, is $(X, S, M_1 \cup M_2)$ compact?
- (2) If (X_i, S_i, M_i) ($i=1, 2$) are compact, is $(X_1 \times X_2, S_1 \times S_2, M_1 \times M_2)$ compact? Here $X_1 \times X_2$ is the product space, $S_1 \times S_2$ the σ -field generated by the S_1 and S_2 cylinder sets and $M_1 \times M_2$ the set of product measures.

We begin with giving the negative answer to the first problem. We consider the following example. Let $X = R^2$ and C be the class of all cones in R^2 . The empty set ϕ is assumed to be a cone. Let S be $\{(E - A) \cup B \mid E \in C; A \text{ and } B \text{ are countable}\}$. Then S becomes a σ -field and it is atomic ($\rho(x) = \{x\}$ for all

x). Let P_x be a probability measure allotting the total mass to a point x and let M_1 be $\{P_x | x \in U = \{(x_1, x_2) | x_1^2 + x_2^2 = 1\}\}$. We shall show that (X, S, M_1) is compact. Let $(f_{P_x})_{x \in U}$ be in the closure of $i_p(B_p(X, S, M_1))$. Define $g(x)$ as follows. If $x \neq 0$, $g(x) = f_{P_{\alpha x}}(\alpha x)$ where α is a real with $\alpha x \in U$ and if $x = 0$, $g(x) = 0$. g is clearly S -measurable and $g = f_{P_x}[P_x]$ for all $P_x \in M_1$. So we have $i_p(g) = (f_{P_x})_{x \in U}$ and therefore $i_p(B_p(X, S, M_1))$ is closed. Hence (X, S, M_1) is compact.

Define a probability measure \tilde{P} as follows. For each $F = (E - A) \cup B (\in S)$, where E is a cone and A, B are countable, $\tilde{P}(F) = 1$ if $(1, 0) \in E$ and $\tilde{P}(F) = 0$ if $(1, 0) \notin E$. Putting $M_2 = \{\tilde{P}\}$, (X, S, M_2) is clearly compact since M_2 is dominated. To prove that $(X, S, M_1 \cup M_2)$ is not compact, we use Theorem 1.3. The set U in this example corresponds to the set H in Theorem 1.3. Since X is an only cone containing U , we have $\Gamma(U | S, \tilde{P}^2) = 1$. By Theorem 1.3 this shows that $(X, S, M_1 \cup M_2)$ is not compact.

Before proceeding to the second problem, we shall prove a lemma. Henceforth G^σ will denote the σ -field generated by G .

Lemma. *If B in G^σ separates x and y , then there exists A in G which separates them.*

Proof. Let J be the set of all such B 's in the statement of our lemma. Then it is easy to show that J is a σ -field and $G \subset J$. Hence we have $G^\sigma = J$ and the proof is completed.

Now we consider the second problem and give the following example. Suppose that $|X| > 2^{\aleph_0}$, where $|X|$ means the cardinality of X . Let M be the set of all discrete probability measures on $(X, 2^X)$. Let X_i, S_i and M_i ($i=1, 2$) be two copies of $X, 2^X$ and M respectively. Then (X_i, S_i, M_i) ($i=1, 2$) are compact due to Theorem 1.1. But we shall show that $(X_1 \times X_2, S_1 \times S_2, M_1 \times M_2)$ is not compact. For this purpose at first we shall prove $\Delta \notin S_1 \times S_2$, where Δ is the diagonal $\{(x, x) | x \in X\}$ of $X_1 \times X_2$. Suppose that Δ is contained in $S_1 \times S_2$. Noting that $\{A \times B | A, B \in 2^X\}$ is a set of generators of $S_1 \times S_2$ we can choose $A_i \times B_i$ ($i=1, 2, \dots$) from it such that $\Delta \in \{A_i \times B_i | i=1, 2, \dots\}^\sigma$ ([9], p. 24). We consider a mapping φ from X into $2^{\{A_1, B_1, A_2, B_2, \dots\}}$ such that $\varphi(x) = \{A_i(B_i) | x \in A_i(B_i)\}$. Since $|X| > 2^{\aleph_0}$ there exist x and y such that $\varphi(x) = \varphi(y)$. Since $(x, x) \in \Delta$, $(x, y) \notin \Delta$, it follows from the above lemma that there exists $A_{i_0} \times B_{i_0}$ which separates (x, x) and (x, y) . So it occurs that B_{i_0} separates x and y , which is contradictory to $\varphi(x) = \varphi(y)$. Hence we have $\Delta \notin S_1 \times S_2$. The facts that every point is a member of $S_1 \times S_2$ and that $\Delta \notin S_1 \times S_2$ imply that $S_1 \times S_2$ is not complete. Hence it follows from Theorem 1.1 that $(X_1 \times X_2, S_1 \times S_2, M_1 \times M_2)$ is not compact. Thus we get the negative answer to the second problem.

WASEDA UNIVERSITY
TOKYO UNIVERSITY OF FISHERIES

References

- [1] P. R. Halmos and L. J. Savage: *Application of the Radon-Nikodym theorem to the theory of sufficient statistics*, Ann. Math. Statist. **20** (1949), 225–241.
- [2] R. R. Bahadur: *Sufficiency and statistical decision functions*, Ann. Math. Statist. **25** (1954), 423–462.
- [3] T. S. Pitcher: *A more general property than domination for sets of probability measures*, Pacific J. Math. **15** (1965), 597–611.
- [4] S. Banach et C. Kuratowski: *Sur une généralisation du problème de la mesure*, Fund. Math. **XIV** (1929), 127–131.
- [5] T.S. Pitcher: *Sets of measures not admitting necessary and sufficient statistics or subfields*, Ann. Math. Statist. **28** (1957), 267–268.
- [6] D.L. Burkholder: *Sufficiency in the undominated case*, Ann. Math. Statist. **32** (1961), 1191–1200.
- [7] D. Basu and J.K. Ghosh: *Sufficient statistics in sampling from a finite universe*, Proceeding of the 36th session of the International Statistical Institute, 1967, 850–859.
- [8] H. Morimoto: *Statistical structure of the problem of sampling from finite populations*, (to appear).
- [9] P.R. Halmos: *Measure Theory*, D. van Nostrand Co., New York, 1950.