PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 15 (1995), pp. 353–363

ON AN EXTENDED NOTION OF COMMON CONDITIONAL PROBABILITY

BY

H. HEYER (TÜBINGEN) AND S. YAMADA (TOKYO)

Abstract. The authors provide a construction of common conditional probabilities given pairwise sufficient σ-fields under the hypothesis that the underlying statistical experiment is majorized. The results are compared with those previously known for more restricted situations and then applied to a new characterization of sufficient σ-fields within the class of pairwise sufficient ones.

1. Introduction. It was J. Neyman who in his pioneering work [8] of 1936 laid the foundations of an analytic approach to what is nowadays called the *theory of sufficient decision procedures.* The so-called *Neyman criterion* put into rigorous measure-theoretic terms in the publication [6] of Halmos and Savage of 1949 has remained a mark of orientation to all later attempts to characterize sufficient statistics, sub-σ-fields, sublattices, projectors, kernels and the like in order to meet the requirements articulated in the applications. While Fisher conveyed only an intuitive understanding of sufficiency in his pertinent paper [4] of 1922, LeCam in [7] provided the ultimate precision; applying Banach lattice theory he achieved independence of his approach from conditions of domination, herewith overcoming the pathologies pointed out by Pitcher [9] and Burkholder [2] on the nonexistence of smallest sufficient statistics and on nonsufficient statistics with finer partitions of the sample space than sufficient statistics, respectively. Along with LeCam's contributions on the subject weaker notions of sufficiency and at the same time replacements for the traditional condition of domination gained increasing interest. In fact, pairwise sufficiency turned out to be the most natural notion and from a slightly different point of view also pairwise sufficiency with supports introduced in [5] by Ghosh, Morimoto and Yamada. The latter authors achieved a generalization of the famous Neyman factorization just in terms of their concept of pairwise sufficiency with supports.

As is well known a σ-field is sufficient if there exists a conditional probability given this σ-field that is common to all measures of the underlying experiment. Using pivotal measures which generalize the classical dominating measures Yamada showed in [13] that there exist common conditional probabilities given pairwise sufficient σ-fields with support. Although his method is based on some lattice-theoretic ideas the spirit of the generalization remains devoted to the work of Halmos and Savage.

In the present paper the authors provide a construction of common conditional probabilities given pairwise sufficient σ-fields under the still weaker hypothesis that the underlying experiment is majorized in the sense of Siebert [10]. The results are compared with those previously known (coincidence of the corresponding common conditional operators, their mutual factorization) and then applied to a new characterization of sufficient σ-fields within the pairwise sufficient ones.

2. Preliminaries. Throughout our exposition we are dealing with classical statistical experiments of the form $E := (X, \mathfrak{A}, \mathscr{P})$, where $\mathscr{P}$ denotes a parametrized family $\{P_\theta\colon \theta \in \Theta\}$ of probability measures P_θ on the measurable space $(X, \mathfrak{A})$. For any sub-σ-field $\mathfrak{B}$ of $\mathfrak{A}$ we consider the subexperiment $E(\mathfrak{B}) = (X, \mathfrak{B}, \mathscr{P}|\mathfrak{B})$ of E with the corresponding family $\mathscr{P}|\mathfrak{B} = \{P_\theta|\mathfrak{B}\colon \theta \in \Theta\}$ of restrictions of P_θ to $\mathfrak{B}$. A sub-σ-field $\mathfrak{B}$ of $\mathfrak{A}$ is called *sufficient* for E or $\mathscr{P}$ if for each $A \in \mathfrak{A}$ there exists a *common conditional probability* $\mathrm{E}(1_A|\mathfrak{B}, \mathscr{P})$ of A given $\mathfrak{B}$ in the sense that

$$\int_B \mathrm{E}(1_A|\mathfrak{B}, \mathscr{P})\,dP_\theta = P_\theta(A \cap B) \quad \text{for all } B \in \mathfrak{B} \text{ and all } \theta \in \Theta.$$

Clearly, $\mathfrak{B}$ is said to be *pairwise sufficient* for $\mathscr{P}$ if $\mathfrak{B}$ is sufficient for all two-element subsets $\mathscr{P}_0$ of $\mathscr{P}$.

In order to characterize pairwise sufficient sub-σ-fields $\mathfrak{B}$ of $\mathfrak{A}$ in terms of LeCam's deficiency [7] we need some tools from the Banach lattice approach to statistical decision theory. For an experiment $E = (X, \mathfrak{A}, \mathscr{P})$ the *L-space* of E is introduced as the band $L(E)$ generated in the space $\mathscr{M}^b(X, \mathfrak{A})$ of all bounded measures on $(X, \mathfrak{A})$ by the set $\mathscr{P}$. The topological dual $M(E)$ of $L(E)$ is called the *M-space* of E. Let $\langle\cdot,\cdot\rangle$ denote the canonical bilinear functional on $L(E) \times M(E)$. Given two experiments $E = (X, \mathfrak{A}, \mathscr{P})$ and $F = (Y, \mathfrak{B}, \mathscr{Q})$, where $\mathscr{Q} := \{Q_\theta\colon \theta \in \Theta\}$ is a subset of probability measures on $(Y, \mathfrak{B})$, we shall consider *transitions* from E to F as positive linear mappings $T\colon L(E) \to L(F)$ satisfying the equality $\|T\mu\| = \|\mu\|$ whenever $\mu \in L(E)_+$. The totality of transitions from E to F will be abbreviated by $\mathrm{Trans}(E, F)$. Now the *deficiency of E with respect to F* is given by

$$\delta(E, F) := \inf_{T \in \mathrm{Trans}(E,F)} \sup_{\theta \in \Theta} \|T(P_\theta) - Q_\theta\|.$$

δ measures the information lost when replacing the experiment E by the experiment F. In fact, E is called *more informative* than F if $\delta(E, F) = 0$. It is a well-known and most applicable fact that a sub-σ-field $\mathfrak{B}$ of $\mathfrak{A}$ is pairwise sufficient for $\mathscr{P}$ iff $\delta(E(\mathfrak{B}), E) = 0$. In order to show this equality it is often useful to exhibit a transition T from $E(\mathfrak{B})$ which renders $\delta(E(\mathfrak{B}), E)$ zero in the sense that $T(P_\theta|\mathfrak{B}) = P_\theta$ for all $\theta \in \Theta$.

Since the pioneering paper of Burkholder [2] we know that sufficiency ought to be studied without domination of the underlying experiments. It was an idea of Siebert [10] to introduce the more general notion of majorization. Given an experiment $E := (X, \mathfrak{A}, \mathscr{P})$ with $\mathscr{P} = \{P_\theta: \theta \in \Theta\}$ and any measure m on $(X, \mathfrak{A})$ we introduce the family $\mathfrak{N}(m) := \{A \in \mathfrak{A}: m(A) = 0\}$ of m-null sets. Accordingly, for any subset $\mathscr{Q}$ of $\mathscr{P}$ we consider the family

$$\mathfrak{N}(\mathscr{Q}) := \{A \in \mathfrak{A}: A \in \mathfrak{N}(P_\theta) \text{ for all } P_\theta \in \mathscr{Q}\}.$$

Finally, the notation $\mathscr{P} \ll m$ stands for the inclusion

$$\mathfrak{N}(m) \subset \bigcap \{\mathfrak{N}(P_\theta): \theta \in \Theta\}.$$

Now, the experiment E is said to be *majorized* if there exists a *majorizing measure* m on $(X, \mathfrak{A})$ in the sense that for each $\theta \in \Theta$ the measure P_θ has an m-density dP_θ/dm. In the case where $\mathscr{P} \sim m$ which means that $\mathfrak{N}(\mathscr{P}) = \mathfrak{N}(m)$ the measure m is called an *equivalent majorizing measure* for E.

Given a majorized experiment E it has been shown by Torgersen [12] that its M-space $M(E)$ consists exactly of all countably coherent families $\{f(\cdot, \theta): \theta \in \Theta\}$ of $\mathfrak{A}$-measurable functions $f(\cdot, \theta)$ on X that are *essentially bounded* in the sense that

$$\sup\{P_\theta\text{-ess}\sup f(\cdot, \theta): \theta \in \Theta\} < \infty.$$

On the other hand, we know that for every countable subfamily $\mathscr{P}_0 := \{P_{\theta_i}: i \geq 1\}$ of $\mathscr{P}$ there exists an $\mathfrak{A}$-measurable function f on X such that $f = f(\cdot, \theta_i)$ P_{θ_i}-a.s. for all $i \geq 1$. Now, let μ be an element of $L(E)$. Then there exists a countable subfamily $\mathscr{P}_0$ of $\mathscr{P}$ such that μ belongs to the band generated by $\mathscr{P}_0$. Writing

$$f(\mathscr{P}) := \{f(\cdot, \theta): \theta \in \Theta\}$$

one can prove that

$$\langle \mu, f(\mathscr{P})\rangle = \int f\, d\mu.$$

The integral (T)-$\int f(\mathscr{P})\,d\mu$ so defined will be called the *(T)-integral of $f(\mathscr{P})$ with respect to μ.*

3. Extended common conditional probabilities for pairwise sufficient sub-σ-fields. Let $E = (X, \mathfrak{A}, \mathscr{P})$ with $\mathscr{P} := \{P_\theta: \theta \in \Theta\}$ be a majorized experiment and let $\mathfrak{B}$ be a sub-σ-field of $\mathfrak{A}$ that is pairwise sufficient for $\mathscr{P}$. Along with E we shall consider the subexperiment $E(\mathfrak{B})$ of E generated by $\mathfrak{B}$. For every $A \in \mathfrak{A}$ we introduce the family

$$f_A(\mathscr{P}) = \{\mathrm{E}(1_A|\mathfrak{B}, P_\theta): \theta \in \Theta\}$$

of conditional probabilities $E(1_A|\mathfrak{B}, P_\theta)$ of 1_A given $\mathfrak{B}$ under P_θ. From the discussion in Section 2 we infer that $f_A(\mathscr{P})$ is an element of the M-space $M(E(\mathfrak{B}))$ of the subexperiment $E(\mathfrak{B})$. Hence for every $\mu \in L(E(\mathfrak{B}))$ the number $\langle \mu, f_A(\mathscr{P})\rangle$ can be calculated as the (T)-integral of $f_A(\mathscr{P})$ with respect to μ given as

$$(T)\text{-}\int f_A(\mathscr{P})d\mu = \int E(1_A|\mathfrak{B}, \mathscr{P}_0)d\mu,$$

where $\mathscr{P}_0$ denotes a countable subfamily of $\mathscr{P}$ such that μ belongs to the band generated in $\mathscr{M}^b(X, \mathfrak{B})$ by $\mathscr{P}_0|\mathfrak{B}$. Since $\mathfrak{B}$ is assumed to be sufficient for $\mathscr{P}_0$, there exists the common conditional expectation $E(1_A|\mathfrak{B}, \mathscr{P}_0)$ defined as usual by

$$\int_B E(1_A|\mathfrak{B}, \mathscr{P}_0)dP = P(A\cap B) \quad \text{for all } P\in\mathscr{P}_0 \text{ and } B\in\mathfrak{B}.$$

In the following we first want to study the mapping $T\colon\ L(E(\mathfrak{B})) \to L(E)$ given by

$$(T\mu)(A) := \langle \mu, f_A(\mathscr{P})\rangle \quad \text{for all } \mu\in L(E(\mathfrak{B})),\ A\in\mathfrak{A},$$

as well as its adjoint $T'\colon M(E) \to M(E(\mathfrak{B}))$.

Clearly, $T(P_\theta|\mathfrak{B}) = P_\theta$ for all $\theta\in\Theta$; hence T is a transition from $L(E(\mathfrak{B}))$ to $L(E)$ rendering the deficiency $\delta(E(\mathfrak{B}), E)$ zero. We note that in general this property does not imply that there exists a Markov kernel T_1 from $(X, \mathfrak{B})$ to $(X, \mathfrak{A})$ such that $T_1(P_\theta|\mathfrak{B}) = P_\theta$ for all $\theta\in\Theta$.

We shall collect some properties of the adjoint T'.

3.1. PROPERTIES OF THE ADJOINT T'. *Let $f(\mathscr{P})\in M(E)$ and $g(\mathscr{P})\in M(E(\mathfrak{B}))$. Then*

3.1.1. $T'(f(\mathscr{P})g(\mathscr{P})) = g(\mathscr{P})T'f(\mathscr{P})$ *(smoothing)*;

3.1.2. $T'^2 f(\mathscr{P}) = T'f(\mathscr{P})$ *(idempotency)*;

3.1.3. $T'g(\mathscr{P}) = g(\mathscr{P})$ *(fixed point property)*.

Subsequently, we shall employ the symbol $\varrho := \varrho_{\mathfrak{A}}$ for the canonical embedding of the space $\mathfrak{M}^b(X, \mathfrak{A})$ of all bounded $\mathfrak{A}$-measurable functions on X into the M-space $M(E)$ of E. Obviously,

$$\varrho(1_A) = \{1_A\colon\ \theta\in\Theta\} \in M(E) \quad \text{for all } A\in\mathfrak{A}.$$

3.2. THEOREM. *Let $A\in\mathfrak{A}$.*

(i) *For all $B\in\mathfrak{B}$ and $\theta\in\Theta$ we have*

$$(T)\text{-}\int_B T'\varrho(1_A)d(I_A P_\theta|\mathfrak{B}) = P_\theta(A\cap B),$$

where

(ii) $T'\varrho(1_A) = \{E(1_A|\mathfrak{B}, P_\theta)\colon\ \theta\in\Theta\} = f_A(\mathscr{P})$.

Proof. (i) For all $A \in \mathfrak{A}$, $B \in \mathfrak{B}$ and $\theta \in \Theta$ we obtain

$$\langle T(1_B P_\theta|\mathfrak{B}), \varrho(1_A)\rangle = \langle 1_B P_\theta|\mathfrak{B}, T'\varrho(1_A)\rangle$$
$$= (T)\text{-}\int T'\varrho(1_A)d(1_B P_\theta|\mathfrak{B}).$$

On the other hand, $1_B P_\theta|\mathfrak{B} \ll P_\theta|\mathfrak{B}$; hence $1_B P_\theta|\mathfrak{B}$ belongs to the band generated by the family $\{P_\theta|\mathfrak{B}: \theta \in \Theta\}$. Thus

$$\langle T(1_B P_\theta|\mathfrak{B}), \varrho(1_A)\rangle = T(1_B P_\theta|\mathfrak{B})(A)$$
$$= \int \mathrm{E}(1_A|\mathfrak{B}, P_\theta)d(1_B P_\theta|\mathfrak{B}) = P_\theta(A \cap B)$$

whenever $A \in \mathfrak{A}$, $B \in \mathfrak{B}$ and $\theta \in \Theta$.

(ii) Suppose now that for $A \in \mathfrak{A}$

$$T'\varrho(1_A) = \{f_A(\cdot, \theta): \theta \in \Theta\}.$$

Then

$$(T)\text{-}\int T'\varrho(1_A)d(1_B P_\theta|\mathfrak{B}) = \int f_A(\cdot, \theta)d(1_B P_\theta|\mathfrak{B})$$
$$= \int_B f_A(\cdot, \theta)d(P_\theta|\mathfrak{B}).$$

But, by the proof of (i),

$$(T)\text{-}\int T'\varrho(1_A)d(1_B P_\theta|\mathfrak{B}) = P_\theta(A \cap B) = \int_B \mathrm{E}(1_A|\mathfrak{B}, P_\theta)d(P_\theta|\mathfrak{B});$$

consequently,

$$f_A(\cdot, \theta) = \mathrm{E}(1_A|\mathfrak{B}, P_\theta) \quad P_\theta\text{-a.s.} \quad \text{for all } \theta \in \Theta. \blacksquare$$

The proof of part (i) of the theorem yields

3.3. COROLLARY. *For all $B \in \mathfrak{B}$ and $\theta \in \Theta$*

$$T(1_B P_\theta|\mathfrak{B}) = 1_B T(P_\theta|\mathfrak{B}).$$

One just applies the fact that $TP_\theta|\mathfrak{B} = P_\theta$ whenever $\theta \in \Theta$.

3.4. DEFINITION. The element $T'\varrho(1_A)$ of $M(E(\mathfrak{B}))$ constructed in Theorem 3.2 is called the *common conditional probability of A given $\mathfrak{B}$ in the extended sense.*

Occasionally, the mapping T' is referred to as the *common conditional probability operator given $\mathfrak{B}$.*

3.5. Remark. In the work of LeCam [7] more general transitions T are considered for which the common conditional probability operator arises as the idempotent pointwise limit of the sequence of finite sums $n^{-1}\sum_{k=1}^{n} T'^k$.

In our case, however, the application of this ergodic type result becomes superfluous since T' is idempotent by Property 3.1.2.

3.6. Remark. It has also been established in LeCam's work that for any transition T: $L(E(\mathfrak{B})) \to L(E)$ which renders $\delta(E(\mathfrak{B}), E)$ zero the dual T' is *expectation invariant* in the sense that $T'(1_B v) = 1_B T'(v)$ whenever $B \in \mathfrak{B}$, $v \in M(E)$.

The transition T that we are concerned with has the weaker property expressed in Corollary 3.3. From this property we may (only) conclude that

$$\langle P_\theta | \mathfrak{B}, T'(1_B v)\rangle = \langle P_\theta | \mathfrak{B}, 1_B T'(v)\rangle$$

for all $\theta \in \Theta$, $B \in \mathfrak{B}$ and $v \in M(E)$, an identity which has been applied in the proof of Theorem 3.2.

We note that under some completion assumption $\mathfrak{B}$ turns out to be the invariant σ-field of T' in the sense that

$$\mathfrak{B} = \{B \in \mathfrak{A}\colon\ T'\varrho(1_B) = \varrho(1_B)\}$$

(see Property 3.1.3).

4. Comparison with the case of pairwise sufficiency with supports. A brief review of the basic notion appearing in the title of the section seems to be in order.

Our standard datum remains an experiment $E = (X, \mathfrak{A}, \mathscr{P})$ with $\mathscr{P} := \{P_\theta\colon\ \theta \in \Theta\}$. For any $P \in \mathscr{P}$ a set $S(P) \in \mathfrak{A}$ such that $P(S(P)) = 1$ and for every $A \in \mathfrak{A}$ with $A \subset S(P)$ and $P(A) = 0$ we have $A \in \mathfrak{N}(\mathscr{P})$ is called a *support of P for E*. If for a sub-σ-field $\mathfrak{B}$ of $\mathfrak{A}$ there exists a support of P belonging to $\mathfrak{B}$ for all $P \in \mathscr{P}$, then $\mathfrak{B}$ is said to *contain supports*. Concerning the existence of supports one has the equivalence of the following statements (Diepenbrock [3]):

(a) Every $P \in \mathscr{P}$ has a support for E.

(b) E is majorized.

(c) There exists an equivalent majorizing measure for E.

The following notions of pairwise sufficiency with support and pivotality have been introduced by Ghosh et al. in [5].

Let E be a majorized experiment.

4.1. Definition. A sub-σ-field $\mathfrak{B}$ of $\mathfrak{A}$ is said to be *pairwise sufficient with supports* (a PSS σ-field) for E if

(i) $\mathfrak{B}$ is pairwise sufficient for $\mathscr{P}$, and

(ii) $\mathfrak{B}$ contains supports.

4.2. Definition. An equivalent majorizing measure n for E is called *pivotal* (for E) if any given sub-σ-field $\mathfrak{B}$ of $\mathfrak{A}$ is PSS iff there exists a $\mathfrak{B}$-measurable density dP_θ/dn for each $\theta \in \Theta$.

It is long known that majorized experiments admit pivotal measures.

One of the most striking applications of the notion of PSS σ-fields is their characterization in terms of a generalized Neyman factorization (see [5]).

In his previous work [13] Yamada discussed in detail a special case of common conditional probability in the extended sense.

Again E is assumed to be majorized. Let $\mathfrak{B}$ be a sub-σ-field of $\mathfrak{A}$. Then for any pivotal measure n for E its restriction $n|\mathfrak{B}$ to $\mathfrak{B}$ is pivotal for the subexperiment $E(\mathfrak{B})$. From the general theory of Banach lattices one infers that

$$L(E) = \{fn: \ f \in L^1(X, \mathfrak{B}, n|\mathfrak{B})\}$$

as well as

$$L(E(\mathfrak{B})) = \{f(n|\mathfrak{B}): \ f \in L^1(X, \mathfrak{B}, n|\mathfrak{B})\}.$$

We define the mapping T_n: $L(E(\mathfrak{B})) \to L(E)$ by

$$T_n(f(n|\mathfrak{B})) := fn \qquad \text{for all } f \in L^1(X, \mathfrak{B}, n|\mathfrak{B}).$$

Then T_n is a transition from $E(\mathfrak{B})$ to E rendering the deficiency $\delta(E(\mathfrak{B}), E)$ zero, and its adjoint T_n' enjoys the idempotency property as well as properties analogous to (i) and (ii) of Theorem 3.2. More precisely, the common conditional probability $T_n'\varrho(1_A)$ of A given $\mathfrak{B}$ with respect to n has the properties collected in the following

4.3. Theorem (Yamada [13], [14]). (i) $(T)\text{-}\int T_n'\varrho(1_A)d(1_B n|\mathfrak{B}) = n(A\cap B)$ *for all* $A \in \mathfrak{A}$, $B \in \mathfrak{B}$ *with* $n(B) < \infty$.

(ii) $(T)\text{-}\int T_n'\varrho(1_A)d(1_B P_\theta|\mathfrak{B}) = P_\theta(A\cap B)$ *for all* $A \in \mathfrak{A}$, $B \in \mathfrak{B}$, *and* $\theta \in \Theta$.

(iii) $T_n'\varrho(1_A) = \{\mathrm{E}(1_A|\mathfrak{B}, P_\theta)\colon \ \theta \in \Theta\}$ *for all* $A \in \mathfrak{A}$.

4.4. Remark. If $\mathfrak{B}$ is sufficient, then by Theorem 4.3

$$T_n'\varrho(1_A) = \mathrm{E}(1_A|\mathfrak{B}, P_\theta) = \mathrm{E}(1_A|\mathfrak{B}, n) \ \ n\text{-a.e.}$$

This conclusion corresponds to the classical case of a dominated experiment E. We may say that in the general situation of a majorized experiment PSS σ-fields are *sufficient with respect to (T)-integrals.*

In fact, the representation (i) of Theorem 4.3 is characteristic of a PSS σ-field within the framework of majorized experiments, as we shall show now.

4.5. Theorem. *Let E be a majorized experiment with equivalent majorizing measure n, and let $\mathfrak{B}$ be a sub-σ-field of $\mathfrak{A}$. Suppose that $\mathfrak{B}$ is PSS. Then the operators T_n and T introduced prior to Theorems* 4.3 *and* 3.2, *respectively, coincide.*

Proof. We fix a set $B \in \mathfrak{B}$ such that $n(B) < \infty$. Next we note that for every $A \in \mathfrak{A}$

$$\begin{aligned}(T)\text{-}\int T_n'\varrho(1_A)d(1_B n|\mathfrak{B}) &= \langle 1_B n|\mathfrak{B}, T_n'\varrho(1_A)\rangle = \langle T_n(1_B n|\mathfrak{B}), \varrho(1_A)\rangle \\ &= \langle 1_B n, \varrho(1_A)\rangle = (T)\text{-}\int \varrho(1_A)d(1_B n).\end{aligned}$$

Moreover, there exists a countable subfamily $\mathscr{P}_0$ of $\mathscr{P}$ such that the measure $1_B n$ belongs to the band generated by $\mathscr{P}_0$. Thus

$$(1)\qquad (T)\text{-}\int \varrho(1_A)d(1_B n) = \int 1_A d(1_B n) = n(A\cap B).$$

On the other hand, we infer from [14] that

$$T_n' \varrho(1_A) = \{\mathrm{E}(1_A|\mathfrak{B}, P_\theta)\colon \theta\in\Theta\};$$

hence

$$(2)\qquad (T)\text{-}\int T_n'\varrho(1_A)d(1_B n|\mathfrak{B}) = \int \mathrm{E}(1_A|\mathfrak{B}, \mathscr{P}_0)d(1_B n|\mathfrak{B}).$$

We conclude that $\mathrm{E}(1_A|\mathfrak{B}, \mathscr{P}_0)$ is a conditional probability of A given $\mathfrak{B}$ common to all $P\in\mathscr{P}_0$.

Now we apply (1) and (2) to obtain

$$\int \mathrm{E}(1_A|\mathfrak{B}, \mathscr{P}_0)d(1_B n|\mathfrak{B}) = n(A\cap B)$$

for all $A\in\mathfrak{A}$ and $B\in\mathfrak{B}$ with $n(B)<\infty$ or; more generally,

$$\int \mathrm{E}(1_A|\mathfrak{B}, \mathscr{P}_0)f d(n|\mathfrak{B}) = \int_A f dn$$

whenever $f\in L^1(X, \mathfrak{B}, n|\mathfrak{B})$. (Here it has to be observed that $\mathscr{P}_0$ depends on $f(n|\mathfrak{B})$.)

But this implies for all $A\in\mathfrak{A}$ and $f\in L^1(X, \mathfrak{B}, n|\mathfrak{B})$ the following chain of equalities

$$T_n(f(n|\mathfrak{B}))(A) = (fn)(A) = \int_A f dn$$

$$= \int \mathrm{E}(1_A|\mathfrak{B}, \mathscr{P}_0)f d(n|\mathfrak{B}) = T(f(n|\mathfrak{B}))(A),$$

i.e. $T_n = T$ on $L(E(\mathfrak{B}))$, which is the desired assertion. ∎

Finally, we discuss the mutual factorization behavior of the mappings T_n and T.

4.6. THEOREM. *Let E be a majorized experiment, and let $\mathfrak{B}$ and $\mathfrak{C}$ be two PSS sub-σ-fields of $\mathfrak{A}$.*

(i) *If*

$$T(L(E(\mathfrak{B})))\subset T_n(L(E(\mathfrak{C}))),$$

then there exists a transition T_1 from $E(\mathfrak{B})$ to $E(\mathfrak{C})$ which renders $\delta(E(\mathfrak{B}), E(\mathfrak{C}))$ zero and satisfies $T_n T_1 = T$. The transition T_1 is uniquely determined by these properties and given by

$$T_1\mu = (T\mu)|\mathfrak{C} \quad \textit{for all } \mu\in L(E(\mathfrak{B})).$$

(ii) *If*

$$T_n L(E(\mathfrak{C}))\subset T(L(E(\mathfrak{B}))),$$

then there exists a transition T_2 from $E(\mathfrak{C})$ to $E(\mathfrak{B})$ which renders $\delta(E(\mathfrak{C}), E(\mathfrak{B}))$ zero and satisfies $TT_2 = T_n$. The transition T_2 is uniquely determined by these properties and given by

$$T_2\mu = (T_n\mu)|\mathfrak{B} \quad \textit{for all } \mu \in L(E(\mathfrak{C})).$$

Proof. It suffices to show that the mappings T and T_n are one-to-one; the desired assertions then follow from the general inverse mapping theorem (see, e.g., [1]).

As for T we suppose that given $\mu_1, \mu_2 \in L(E(\mathfrak{B}))$ we have $T\mu_1 = T\mu_2$. Then

$$\langle \mu_1 - \mu_2, f_A(\mathscr{P})\rangle = 0 \quad \text{for all } A \in \mathfrak{A}.$$

Thus, for all $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$ we obtain

$$0 = \langle \mu_1 - \mu_2, f_{A\cap B}(\mathscr{P})\rangle = \int \mathrm{E}(1_{A\cap B}|\mathfrak{B}, \mathscr{P}_0)d(\mu_1 - \mu_2)$$
$$= \int_B \mathrm{E}(1_A|\mathfrak{B}, \mathscr{P}_0)d(\mu_1 - \mu_2),$$

where $\mathscr{P}_0$ denotes a countable subfamily of $\mathscr{P}$ such that $\mu_1 - \mu_2$ belongs to the band generated (in $\mathscr{M}^b(X, \mathfrak{A})$) by $\mathscr{P}_0$, and

$$\mathrm{E}(1_{A\cap B}|\mathfrak{B}, \mathscr{P}_0) = 1_B \mathrm{E}(1_A|\mathfrak{B}, \mathscr{P}_0)$$

holds $(\mu_1 - \mu_2)$-a.e. Choosing $A := X$ in the above chain of equalities we see that $\mu_1 = \mu_2$.

As for the mapping T_n we assume that $T_n(f_1(n|\mathfrak{B})) = T_n(f_2(n|\mathfrak{B}))$ for $f_1, f_2 \in L^1(X, \mathfrak{B}, n|\mathfrak{B})$, which means that $f_1 n = f_2 n$ on $\mathfrak{A}$. But this implies that

$$f_1(n|\mathfrak{B}) = (f_1 n)|\mathfrak{B} = (f_2 n)|\mathfrak{B} = f_2(n|\mathfrak{B}),$$

which is the assertion. ∎

4.7. Corollary. *For every $\mu \in L(E(\mathfrak{B}))$ we have $T_n(T\mu|\mathfrak{B}) = T$.*

5. Another characterization of sufficiency (in terms of the extended common conditional probability operator). Again we start with an experiment $E = (X, \mathfrak{A}, \mathscr{P})$ and a sub-σ-field $\mathfrak{B}$ of $\mathfrak{A}$. By $\varrho_{\mathfrak{B}}$ we denote the canonical embedding from $\mathfrak{M}^b(X, \mathfrak{B})$ into $M(E(\mathfrak{B}))$.

5.1. Theorem. *Suppose that the σ-field $\mathfrak{B}$ is pairwise sufficient for $\mathscr{P}$ so that the extended common conditional operator T' given $\mathfrak{B}$ is defined. Then the following statements are equivalent:*

(i) *$\mathfrak{B}$ is sufficient for $\mathscr{P}$;*

(ii) *$T'\varrho(\mathfrak{M}^b(X, \mathfrak{A})) \subset \varrho_{\mathfrak{B}}(\mathfrak{M}^b(X, \mathfrak{B}))$.*

Proof. Only the implication (ii) ⇒ (i) has to be shown.

Fixing $A \in \mathfrak{A}$ we write

$$T'\varrho(1_A) = \varrho_{\mathfrak{B}}(g_A)$$

with a bounded $\mathfrak{B}$-measurable function g_A on $(X, \mathfrak{A})$. For any $\theta \in \Theta$ and $B \in \mathfrak{B}$ we define the measure $P_\theta|B$ on $(X, \mathfrak{A})$ by

$$P_\theta|B(\cdot) := P_\theta(\cdot \cap B).$$

Then $P_\theta|B \in L(E)$, and hence $(P_\theta|B)|\mathfrak{B} \in L(E(\mathfrak{B}))$.

Now we have

$$T[(P_\theta|B)|\mathfrak{B}](A) = \int \mathrm{E}(1_A|\mathfrak{B}, P_\theta)d[(P_\theta|B)|\mathfrak{B}] = \int_B \mathrm{E}(1_A|\mathfrak{B}, P_\theta)dP_\theta = P_\theta(A \cap B)$$

whenever $B \in \mathfrak{B}$, $\theta \in \Theta$. This implies that

$$\langle (P_\theta|B)|\mathfrak{B}, T'\varrho(1_A)\rangle = \langle T[(P_\theta|B)|\mathfrak{B}], \varrho(1_A)\rangle = \langle P_\theta(\cdot \cap B), \varrho(1_A)\rangle = P_\theta(A \cap B)$$

for all $B \in \mathfrak{B}$, $\theta \in \Theta$. On the other hand, we have

$$\langle (P_\theta|B)|\mathfrak{B}, T'\varrho(1_A)\rangle = \langle (P_\theta|B)|\mathfrak{B}, \varrho_{\mathfrak{B}}(g_A)\rangle = \int g_A d[(P_\theta|B)|\mathfrak{B}] = \int_B g_A d(P_\theta|\mathfrak{B})$$

for all $B \in \mathfrak{B}$, $\theta \in \Theta$. But this implies that

$$\int_B g_A d(P_\theta|\mathfrak{B}) = P_\theta(A \cap B) \quad \text{for all } B \in \mathfrak{B},\ \theta \in \Theta,$$

or equivalently

$$\mathrm{E}(1_A|\mathfrak{B}, P_\theta) = g_A \ \ P_\theta\text{-a.s.} \quad \text{for all } \theta \in \Theta.$$

That is to say, g_A is a common conditional probability of A given $\mathfrak{B}$ with respect to $\mathscr{P}$. The sufficiency of $\mathfrak{B}$ for $\mathscr{P}$ has thus been established. ■

REFERENCES

[1] N. Bourbaki, *Elements of Mathematics, Theory of Sets*, Addison–Wesley, Reading, MA, 1968 (by Hermann).

[2] D. L. Burkholder, *Sufficiency in the undominated case*, Ann. Math. Statist. 32 (1961), pp. 1191–1200.

[3] F. R. Diepenbrock, *Charakterisierung einer allgemeineren Bedingung als Dominiertheit mit Hilfe von lokalisierten Maßen*, Ph. D. Thesis, University of Münster 1971.

[4] R. A. Fisher, *On the mathematical foundation of theoretical statistics*, Philos. Trans. Roy. Soc. London Ser. A 222 (1922), pp. 309–368.

[5] J. K. Ghosh, H. Morimoto and S. Yamada, *Neyman factorization and minimality of pairwise sufficient subfields*, Ann. Statist. 9 (1981), pp. 514–530.

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

[7] L. LeCam, *Sufficiency and approximate sufficiency*, ibidem 35 (1964), pp. 1419–1455.
[8] J. Neyman, *Sur un teorema concernente le considette statistiche sufficienti*, Giorn. Ist. Ital. Att. 6 (1935), pp. 320–334.
[9] T. S. Pitcher, *Sets of measure not admitting necessary and sufficient statistics or subfields*, Ann. Math. Statist. 28 (1957), pp. 267–268.
[10] E. Siebert, *Pairwise sufficiency*, Z. Wahrscheinlichkeitstheorie verw. Gebiete 46 (1979), pp. 237–246.
[11] H. Strasser, *Mathematical Theory of Statistics*, Walter de Gruyter, 1985.
[12] E. N. Torgersen, *Comparison of Statistical Experiments*, Cambridge University Press, 1991.
[13] S. Yamada, *M-space of majorized experiment and pivotal measure*, Statist. Decisions 6 (1988), pp. 163–174.
[14] — *Applications of an extended notion of common conditional probability*, Report of Tokyo Univ. Fish. 26 (1991), pp. 79–87.

Mathematisches Institut der Universität
Auf der Morgenstelle 10 Block C
72076 Tübingen, Germany

Tokyo University of Fisheries
5–7 Konan 4, Minato-ku
Tokyo 108, Japan

Received on 8.2.1994