The Annals of Statistics
1977, Vol. 5, No. 1, 44-64

MIXTURES AND PRODUCTS OF DOMINATED EXPERIMENTS

BY ERIK N. TORGERSEN

University of Oslo

It is shown, using a theorem of Choquet, that any separable experiment is a mixture of experiments admitting boundedly complete and sufficient statistics. The experiments possessing these properties are precisely the experiments which are extremal with respect to mixtures.

Dominated models for independent observations $X_1, \cdots, X_n$ admitting boundedly (or L_p) complete and sufficient statistics, are considered. It is shown that a subset—say $X_1, \cdots, X_m$ where $m < n$—has the same property provided a regularity condition is satisfied. This condition is automatically satisfied when the observations are identically distributed. In the bounded complete case the proof uses the fact that products of experiments are distributive w.r.t. mixtures. More involved arguments are needed for L_p completeness.

1. Introduction, basic facts and notations. Mixtures of experiments with the same finite parameter set were treated by Birnbaum [3, 4] and by the author [18]. Some of these results are here generalized to the case of dominated experiments. We will also present results which, when restricted to experiments with a finite parameter set, appear new.

Some basic facts on comparison of experiments are recapitulated below. It is only the material up to and including Corollary 1.2 which is needed for the logical development in Sections 2 and 3. The remaining part of this section is a short exposition of important and relevant results. Most of the results in this section are, usually in a somewhat different shape, established in Le Cam [9]. The reader may consult Le Cam [9, 12], Heyer [7] and Torgersen [18, 20] for more thorough expositions and for detailed proofs.

An *experiment* will here be defined as a pair $\mathscr{E} = ((\chi, \mathscr{A}); \mu_\theta : \theta \in \Theta)$ where $(\chi, \mathscr{A})$ is a measurable space, the *sample space*, and $(\mu_\theta : \theta \in \Theta)$ is a family of probability measures on $\mathscr{A}$. The set Θ is the *parameter set* of $\mathscr{E}$. The notation of the sample space will often be suppressed. Thus we may write $\mathscr{E} = (\mu_\theta; \theta \in \Theta)$ instead of $\mathscr{E} = ((\chi, \mathscr{A}); \mu_\theta : \theta \in \Theta)$. The *restriction* $(\mu_\theta : \theta \in \Theta_0)$ where Θ_0 is some nonempty subset of Θ will be denoted by $\mathscr{E}_{\Theta_0}$. If $\mathscr{E}_t = ((\chi_t, \mathscr{A}_t), \mu_{\theta t} : \theta \in \Theta)$; $t \in T$ is a family of experiments then their *product* $\prod_T \mathscr{E}_t$ is the experiment $((\chi, \mathscr{A}), \mu_\theta : \theta \in \Theta)$ where $(\chi, \mathscr{A}) = \prod_T (\chi_t, \mathscr{A}_t)$ and $\mu_\theta = \prod_t \mu_{\theta t}$; $\theta \in \Theta$.

The *total variation* of any finite measure μ will be denoted by $||\mu||$. The *total variation distance* between two finite measures μ' and μ'' is $||\mu' - \mu''||$. Metrical notions in this paper are, if not otherwise stated, w.r.t. this metric. If

Received June 1975; revised May 1976.
AMS 1970 *subject classifications*. Primary 62B15; Secondary 62B05.
Key words and phrases. Extremal experiments, isometry criterion.

$\mathscr{E} = (\mu_\theta : \theta \in \Theta)$ is an experiment then we shall denote by $B(\mathscr{E})$ the closure of the linear space spanned by $\{\mu_\theta : \theta \in \Theta\}$.

All experiments considered in this paper will, unless otherwise indicated, have the same parameter set Θ. Θ may be any nonempty set. An experiment $\mathscr{E} = (\mu_\theta : \theta \in \Theta)$ is called *dominated* if the set $\{\mu_\theta : \theta \in \Theta\}$ of probability measures is dominated. We will throughout this section and Section 3 assume that all experiments under consideration are dominated. The set of finite measures μ such that $\mu(A) = 0$ whenever $\mu_\theta(A) \equiv_\theta 0$ will be denoted by $L(\mathscr{E})$. It is shown by Halmos and Savage [6] that a dominated family $(\mu_\theta : \theta \in \Theta)$ of probability measure is dominated by probability measures π of the form $\pi = \sum c_\theta \mu_\theta$ where c is a prior distribution with countable support. If $\mathscr{E} = (\mu_\theta : \theta \in \Theta)$ is a dominated experiment and π is a dominating measure of this form then π will be called an *L-measure* for $\mathscr{E}$. If π is an L-measure for $\mathscr{E}$ then $L(\mathscr{E})$ consists of all finite measures μ which are absolutely continuous w.r.t. π. It was shown by Bahadur [1] that the σ-algebra induced by $(d\mu_\theta/d\pi : \theta \in \Theta)$ is minimal sufficient when π is an L-measure for $\mathscr{E}$.

Let $\mathscr{E} = (\chi, \mathscr{A}, \mu_\theta : \theta \in \Theta)$ be an experiment with L-measure π and let $\mathscr{B}$ be a minimal sufficient sub-σ-algebra of $\mathscr{A}$. Then the set, $V(\mathscr{E})$, of measures μ in $L(\mathscr{E})$ whose Radon–Nikodym derivatives w.r.t. π may be specified $\mathscr{B}$-measurable does not depend on the choice of π nor on the specification of $\mathscr{B}$.

The sets $L(\mathscr{E})$ and $V(\mathscr{E})$ are, according to the terminology in Le Cam [9], respectively the *L-space* of $\mathscr{E}$ and the *minimal equivalent form* of $\mathscr{E}$.

Let $\mathscr{E} = (\mu_\theta : \theta \in \Theta)$ and $\mathscr{F} = (\nu_\theta : \theta \in \Theta)$ be two experiments and let ε be a nonnegative function on Θ. Following Le Cam [9] we shall say that $\mathscr{E}$ is ε-*deficient* w.r.t. $\mathscr{F}$ if for any finite decision space, any bounded loss function and any risk function $\tilde{r}$ obtainable in $\mathscr{F}$, there is a risk function r obtainable in $\mathscr{E}$ so that $r \leqq \tilde{r} + \varepsilon$.

A very interesting criterion for ε-deficiency is Le Cam's randomization criterion (Theorem 3 in [9]). As this criterion is not needed here we shall not elaborate further on it.

It follows from weak compactness ([15], page 118) that $\mathscr{E}$ is ε-deficient w.r.t. $\mathscr{F}$ if and only if $\mathscr{E}_F$ is ε/F deficient w.r.t. $\mathscr{F}_F$ for each finite and nonempty subset F of Θ.

The experiment $\mathscr{E}$ is *more informative than* $\mathscr{F}$ if $\mathscr{E}$ is 0-deficient w.r.t. $\mathscr{F}$. If $\mathscr{E}$ is more informative than $\mathscr{F}$ and $\mathscr{F}$ is more informative than $\mathscr{E}$ then we shall say that $\mathscr{E}$ and $\mathscr{F}$ are *equivalent* and write $\mathscr{E} \sim \mathscr{F}$.

Two useful criterions for equivalence are:

THEOREM 1.1 (Le Cam [9, 12]). *Let $\mathscr{E} = (\mu_\theta : \theta \in \Theta)$ and $\mathscr{F} = (\nu_\theta : \theta \in \Theta)$ be dominated experiments. Let c be a nonnegative function on Θ so that $\mu = \sum_\theta c_\theta \mu_\theta$ and $\nu = \sum_\theta c_\theta \nu_\theta$ are L-measures for, respectively, $\mathscr{E}$ and $\mathscr{F}$. Put $f = (d\mu_\theta/d\mu : \theta \in \Theta)$ and $g = (d\nu_\theta/d\nu : \theta \in \Theta)$. Then the following conditions are equivalent*:

(i) $\mathscr{E} \sim \mathscr{F}$;

(ii) *The linear space spanned by* $\{\mu_\theta : \theta \in \Theta\}$ *is isometric, for the total variation norm, to the linear space spanned by* $\{\nu_\theta : \theta \in \Theta\}$ *by an isometry (necessarily unique) mapping* μ_θ *into* ν_θ;

(iii) $\mu f^{-1} = \nu g^{-1}$.

REMARK. (ii) may also be formulated:

(ii′) $B(\mathscr{E})$ and $B(\mathscr{F})$ are isometric by an isometry, necessarily unique, mapping μ_θ into ν_θ.

PROOF.

(i) $\Rightarrow$ (ii): We may, without loss of generality, assume that Θ is finite. Let $a \in R^\Theta$. Consider the set $\{0, 1\}$ as a decision space and let the loss function L be given by:

$$L_\theta(t) = (-1)^t a_\theta ; \qquad t = 0, 1; \theta \in \Theta .$$

Then the minimum Bayes risks for the uniform prior distribution are, respectively, $-\|\sum_\theta a_\theta \mu_\theta\| m^{-1}$ and $-\|\sum_\theta a_\theta \nu_\theta\| m^{-1}$ where $m = \#\Theta$. Hence

$$\|\textstyle\sum_\theta a_\theta \mu_\theta\| = \|\sum_\theta a_\theta \nu_\theta\| ; \qquad a \in R^\Theta$$

so that (ii) holds.

(ii) $\Rightarrow$ (iii): Let $\theta_0 \in \Theta$, let F be a finite subset of Θ containing θ_0 and let $a \in R^F$. By (ii):

$$\begin{aligned}\textstyle\int (\sum_F a_\theta f_\theta)^+ \, d\mu &= \tfrac{1}{2} \textstyle\int [\sum_F a_\theta f_\theta + |\sum_F a_\theta f_\theta|] \, d\mu \\ &= \tfrac{1}{2} \textstyle\sum_F a_\theta + \|\sum_F a_\theta \mu_\theta\| = \tfrac{1}{2} \sum_F a_\theta + \|\sum_F a_\theta \nu_\theta\| \\ &= \textstyle\int (\sum_F a_\theta g_\theta)^+ \, d\nu .\end{aligned}$$

Differentiating from the right w.r.t. a_{θ_0} we find

$$\mu_{\theta_0}(\textstyle\sum_F a_\theta f_\theta \geqq 0) = \nu_{\theta_0}(\sum_F a_\theta g_\theta \geqq 0) .$$

Let z be a real number such that $\mu(\sum_F a_\theta f_\theta = z) = \nu(\sum_F a_\theta g_\theta = z) = 0$. Put:

$$\begin{aligned}\tilde{a}_\theta &= a_\theta - z c_\theta \qquad \text{when} \quad \theta \in F \\ &= -z c_\theta \qquad\quad \text{when} \quad \theta \notin F .\end{aligned}$$

Let F_n; $n = 1, 2, \cdots$ be an increasing sequence of finite sets such that $\bigcup_n F_n = F \cup \{\theta : c_\theta > 0\}$. Utilizing that $\sum_\theta c_\theta f_\theta = 1$ a.e. μ and that $\sum_\theta c_\theta g_\theta = 1$ a.e. ν we get

$$\lim_{n\to\infty} \textstyle\sum_{F_n} \tilde{a}_\theta f_\theta = \sum_F a_\theta f_\theta - z \qquad \text{a.e.} \quad \mu$$

and

$$\lim_{n\to\infty} \textstyle\sum_{F_n} \tilde{a}_\theta g_\theta = \sum_F a_\theta g_\theta - z \qquad \text{a.e.} \quad \nu .$$

Hence

$$\begin{aligned}\mu_{\theta_0}(\textstyle\sum_F a_\theta f_\theta \geqq z) &= \lim_n \mu_{\theta_0}(\sum_{F_n} \tilde{a}_\theta f_\theta \geqq 0) \\ &= \lim_n \nu_{\theta_0}(\textstyle\sum_{F_n} \tilde{a}_\theta g_\theta \geqq 0) = \nu_{\theta_0}(\sum_F a_\theta g_\theta \geqq z) .\end{aligned}$$

Thus $\mu_{\theta_0}(f \mid F)^{-1}$ and $\nu_{\theta_0}(f \mid F)^{-1}$ coincides on the class of half spaces of R^F. It follows that these measures have the same characteristic functions so that

$\mu_{\theta_0}(f \mid F)^{-1} = \nu_{\theta_0}(g \mid F)^{-1}$. Hence, as this holds for any finite subset F containing θ_0,

$$\mu_\theta f^{-1} = \nu_\theta g^{-1}\,; \qquad \theta \in \Theta$$

so that

$$\mu f^{-1} = \sum c_\theta(\mu_\theta f^{-1}) = \sum c_\theta(\nu_\theta g^{-1}) = \nu g^{-1}\,.$$

(iii) $\Rightarrow$ (i): Suppose $\mu f^{-1} = \nu g^{-1}$ and consider the set $T = \{1, 2, \cdots, k\}$ as a decision space. Let $\tau(t \mid \theta)$; $t \in \Theta$ be the operational characteristic (performance function) of a decision rule ρ in $\mathscr{E}$. We may, since f is sufficient in $\mathscr{E}$, assume that there is a measurable function h from R^Θ to the set of probability distributions on $\{1, 2, \cdots, k\}$ so that $\rho = hf$. Then hg is a decision rule in $\mathscr{F}$ and

$$\tau(t \mid \theta) = \int h_t(f) f_\theta \, d\mu = \int h_t(g) g_\theta \, d\nu\,; \qquad t \in T;\ \theta \in \Theta\,.$$

Thus τ is also obtainable in $\mathscr{F}$. It follows in particular that, for any given loss function, $\mathscr{E}$ and $\mathscr{F}$ determine the same set of risk functions. □

COROLLARY 1.2. *Suppose $\mathscr{E}$ admits a boundedly complete and sufficient sub-σ-algebra and that $\mathscr{E} \sim \mathscr{F}$. Then $\mathscr{F}$ admits a boundedly complete and sufficient sub-σ-algebra.*

PROOF. Let us use the notations of the theorem for the experiments $\mathscr{E}$ and $\mathscr{F}$. By assumption and by Bahadur [1], the σ-algebra induced by f is sufficient and boundedly complete. Hence, by the theorem, the σ-algebra induced by g is sufficient and boundedly complete in $\mathscr{F}$. □

The remainder of this section is a description of a few important and known results. This material is included for completeness and is not needed for the logical development in the next sections.

It was claimed above that $L(\mathscr{E})$ and $V(\mathscr{E})$ was, according to the terminology in Le Cam's paper [9], respectively the L-space of $\mathscr{E}$ and the minimal equivalent form of $\mathscr{E}$. In order to verify this we must check that:

$$L(\mathscr{E}) = \text{the band generated by } \{\mu_\theta : \theta \in \Theta\}$$

and

$$V(\mathscr{E}) = \text{the Banach lattice generated by } \{\mu_\theta : \theta \in \Theta\}\,.$$

The set $\mathscr{M}(\mathscr{E})$ of finite measures on $\mathscr{E}$'s sample space is an ordered linear space for setwise definitions of ordering as well as of linear combinations. With the total variation norm $\mathscr{M}(\mathscr{E})$ becomes a normed linear space. It may then, see [15], be shown that $\mathscr{M}(\mathscr{E})$ with these structures is a Banach lattice which has the additional property that the norm is additive on the cone of nonnegative elements. A Banach lattice with this property is called an L-space. A subset U of $\mathscr{M}(\mathscr{E})$ is called a band if:

(i) $\mu \in U,\ \nu \in \mathscr{M}(\mathscr{E}),\ |\nu| \leqq |\mu| \Rightarrow \nu \in U$

and

(ii) $U_0 \subseteq U,\ \mu = \sup\{\nu : \nu \in U_0\} \in \mathscr{M}(\mathscr{E}) \Rightarrow \mu \in U$.

Let $L_0(\mathscr{E})$ be the set of all measures $\mu \in \mathscr{M}(\mathscr{E})$ such that $-\sum_{i=1}^n \mu_{\theta_i} \leqq \mu \leqq \sum_{i=1}^n \mu_{\theta_i}$ for some finite sequence $(\theta_1, \theta_2, \cdots, \theta_n)$ in Θ. Define also $V_0(\mathscr{E})$ as the set of all measures of the form $\bigvee_{i=1}^n \sum_F a_\theta^{(i)}\mu_\theta - \bigvee_{i=1}^n \sum_F b_\theta^{(i)}\mu_\theta$ for some finite subset F of Θ and for some real constants $a_\theta^{(i)}$, $b_\theta^{(i)}$. The claims above may then be verified by showing that $L_0(\mathscr{E})$ and $V_0(\mathscr{E})$ are dense subvector lattices of, respectively, $L(\mathscr{E})$ and $V(\mathscr{E})$.

By the result of Bahadur cited above any measure μ_θ is in $V(\mathscr{E})$ and it follows directly from the definition of $V(\mathscr{E})$ that $V(\mathscr{E}) \subseteqq L(\mathscr{E})$.It is shown in [9] that $V(\mathscr{E}) = L(\mathscr{E})$ if and only if $\mathscr{A}$ is minimal sufficient. With our definitions of the spaces $V(\mathscr{E})$ and $L(\mathscr{E})$ this follows from the essential uniqueness of Radon–Nikodym derivatives. The author learned from Le Cam several years ago the fact and the proof given below, that $B(\mathscr{E}) = V(\mathscr{E})$ if and only if $\mathscr{A}$ admits a boundedly complete and sufficient sub-σ-algebra. As the experiments having the last property will play an important role in our investigation we formulate these results as:

THEOREM 1.3. *Let $\mathscr{E} = (\chi, \mathscr{A}, \mu_\theta : \theta \in \Theta)$ be a dominated experiment. Then*:

(i) $B(\mathscr{E}) \subseteqq V(\mathscr{E}) \subseteqq L(\mathscr{E})$;

(ii) $B(\mathscr{E}) = V(\mathscr{E})$ *if and only if* $\mathscr{A}$ *admits a boundedly complete and sufficient sub-σ-algebra*;

(iii) $V(\mathscr{E}) = L(\mathscr{E})$ *if and only if* $\mathscr{A}$ *is minimal sufficient*;

(iv) $B(\mathscr{E}) = L(\mathscr{E})$ *if and only if* $\mathscr{A}$ *is boundedly complete.*

PROOF. It suffices, as (i) and (iii) were argued above and (iv) is a consequence of (ii) and (iii), to prove (ii). Let $\tau \in V(\mathscr{E}) - B(\mathscr{E})$ and let π be an L-measure for $\mathscr{E}$. Then there is a continuous linear functional, h, on $L(\mathscr{E})$, i.e., an element of $L(\mathscr{E})^* = L_\infty(\pi)$, strongly separating τ from $B(\mathscr{E})$. We may, since $B(\mathscr{E})$ is linear, assume that $\tau(h) > 0 = \lambda(h)$ when $\lambda \in B(\mathscr{E})$. Let $\mathscr{B}$ be any minimal sufficient sub-σ-algebra of $\mathscr{A}$. Then there is a bounded $\mathscr{B}$-measurable function $\tilde{h}$ on χ so that $\lambda(h) = \lambda(\tilde{h})$ when $\lambda \in V(\mathscr{E})$. It follows that we may as well assume that h is $\mathscr{B}$-measurable. Hence $\mathscr{B}$ cannot, since $\mu_\theta(h) \equiv_\theta 0$ while $\tau(h) > 0$, be boundedly complete. If, conversely, $\mathscr{B}$ is not boundedly complete then there is a bounded $\mathscr{B}$-measurable function h so that $\int |h| \, d\pi > 0$ and $\int h \, d\mu_\theta \equiv_\theta 0$. Then h defines a nonzero linear functional on $V(\mathscr{E})$ which is vanishing on $B(\mathscr{E})$ so that $B(\mathscr{E}) \subset V(\mathscr{E})$.

By Le Cam's isometry criterion, $\mathscr{E} \sim \mathscr{F}$ if and only if $B(\mathscr{E})$ and $B(\mathscr{F})$ are isometric by an isometry mapping μ_θ into ν_θ. It is shown in [9], and it follows also readily from Theorem 1.1, that such an isometry may be extended to an isometric lattice isomorphism between $V(\mathscr{E})$ and $V(\mathscr{F})$. This, together with the completeness criterion described above, implies again that the property of having a boundedly complete and sufficient sub-σ-algebra respects equivalence.

If $\mathscr{E}$ has this property and we want to find out whether $\mathscr{E}$ is more informative than an experiment $\mathscr{F}$ then the following criterion, proved by De Groot in [5], is useful:

THEOREM 1.4. *Let* $\mathscr{E} = (\chi, \mathscr{A}, \mu_\theta : \theta \in \Theta)$ *be an experiment such that* $\mathscr{E}$ *admits a boundedly complete and sufficient sub-*σ*-algebra. Then an experiment* $\mathscr{F}$ *is less informative than* $\mathscr{E}$ *if and only if any power function obtainable in* $\mathscr{F}$ *is obtainable in* $\mathscr{E}$.

REMARK. By [18] and weak compactness this amounts to the fact that $\mathscr{E}$ is more informative than $\mathscr{F}$ if and only if $\mathscr{E}$ is more informative than $\mathscr{F}$ for testing problems.

PROOF. We may, without loss of generality, assume that $\mathscr{A}$ is minimal sufficient and, consequently, boundedly complete. Put $\mathscr{F} = (\mathscr{Y}, \mathscr{B}, \nu_\theta : \theta \in \Theta)$. Let $T = (1, 2, \cdots, k)$ be a finite decision space and let σ be a decision rule in $\mathscr{F}$. We may represent σ as $\sigma = (\sigma_1, \sigma_2, \cdots, \sigma_k)$ where $\sigma_t(y)$ is the probability of deciding t when y is observed. By assumption there are test functions $\rho_1, \cdots, \rho_k$ in $\mathscr{E}$ so that $\int \rho_t \, d\mu_\theta \equiv_\theta \int \sigma_t \, d\nu_\theta$. Then $\int \sum_t \rho_t \, d\mu_\theta \equiv_\theta \int \sum_t \sigma_t \, d\nu_\theta = 1$ so that $\sum_t \rho_t = 1$ a.e. μ_θ for all θ. Redefining ρ on the null set $\{x : \sum_t \rho_t(x) \neq 1\}$ so that $\sum_t \rho_t(x) \equiv_x 1$ we obtain a decision rule in $\mathscr{E}$ having the same operational characteristic as σ has in $\mathscr{F}$. $\square$

2. Convex combinations of dominated experiments and extremal experiments. It was shown in [18] that—in the case of a finite parameter set—an experiment does not have a sufficient and complete sub-σ-algebra if and only if it is equivalent with a proper mixture of two nonequivalent experiments. In trying to generalize results valid for finite Θ we must keep in mind that, in general, we have to distinguish between several types of completeness, In particular bounded completeness does not—see Lehmann ([14], page 152)—imply completeness. It turns out that many of the results in [18] carry over to the infinite case provided we replace "completeness" with "bounded completeness." A more detailed examination of some possible generalizations of the completeness concept for finite Θ will be given in the next section.

The assumption that the measures defining an experiment are all probability measures is not essential for the development below. In fact, most of the results have ([19]) straightforward generalizations to arbitrary families of finite measures.

Consider two experiments $\mathscr{E} = ((\chi, \mathscr{A}), \mu_\theta : \theta \in \Theta)$ and $\mathscr{F} = ((\mathscr{Y}, \mathscr{B}), \nu_\theta : \theta \in \Theta)$ and a number $\tau \in [0, 1]$. A new experiment $\mathscr{G} = ((\mathscr{X}, \mathscr{C}), \sigma_\theta : \theta \in \Theta)$ may then be defined by:

$$\mathscr{X} = [\chi \times \{1\}] \cup [\mathscr{Y} \times \{2\}]$$

$\mathscr{C}$ = the σ-algebra of all subsets of $\mathscr{X}$ of the form

$$[A \times \{1\}] \cup [B \times \{2\}] \quad \text{where} \quad A \in \mathscr{A} \quad \text{and} \quad B \in \mathscr{B}$$

$$\sigma_\theta([A \times \{1\}] \cup [B \times \{2\}]) = (1 - \tau)\mu_\theta(A) + \tau\nu_\theta(B) ; \qquad A \in \mathscr{A}, \quad B \in \mathscr{B}.$$

The notation $(1 - \tau)\mathscr{E} + \tau\mathscr{F}$ will occasionally be used for this experiment.

When $\mathscr{E}$ and $\mathscr{F}$ are dominated, then $(1 - \tau)\mathscr{E} + \tau\mathscr{F}$ is dominated. This and other facts are collected in:

PROPOSITION 2.1. *Let $\mathscr{E}$, $\mathscr{F}$ and τ be as above, let c be a nonnegative function on Θ such that $\sum_\theta c_\theta \mu_\theta$ and $\sum c_\theta \nu_\theta$ are L-measures for respectively $\mathscr{E}$ and $\mathscr{F}$. Then $\sum_\theta c_\theta \sigma_\theta$ is an L-measure for $(1-\tau)\mathscr{E} + \tau\mathscr{F}$. Put*

$$\mu = \textstyle\sum_\theta c_\theta \mu_\theta\,, \qquad \nu = \sum_\theta c_\theta \nu_\theta\,, \qquad \sigma = \sum_\theta c_\theta \sigma_\theta\,,$$
$$f_\theta = d\mu_{\theta/d\mu}\,, \qquad g_\theta = d\nu_{\theta/d\nu}\,, \qquad h_\theta = d\sigma_{\theta/d\sigma}\,,$$
$$f = (f_\theta : \theta \in \Theta)\,, \qquad g = (g_\theta : \theta \in \Theta)\,, \qquad h = (h_\theta : \theta \in \Theta)\,.$$

Then

$$\sigma([A \times \{1\}] \cup [B \times \{2\}]) = (1-\tau)\mu(A) + \tau\nu(B)$$

and

$$\sigma h^{-1} = (1-\tau)\mu f^{-1} + \tau\nu g^{-1}\,.$$

One possible specification of h_θ is:

$$h_\theta((x, 1)) = f_\theta(x)\,, \quad x \in \chi$$
$$h_\theta((y, 2)) = g_\theta(y)\,; \quad y \in \mathscr{Y}\,.$$

PROOF. Note first that $\sigma([A \times \{1\}] \cup [B \times \{2\}]) = (1-\tau)\mu(A) + \tau\nu(B)$. Hence σ is an L-measure for $(1-\tau)\mathscr{E} + \tau\mathscr{F}$.

It is a matter of checking that $d\sigma_{\theta/d\sigma}$ may be specified as stated. Finally let κ be any bounded measurable function on R^Θ. Then $\sigma(\kappa \circ h) = \sigma((\kappa \circ h)I_{\chi\times\{1\}} + (\kappa \circ h)I_{\mathscr{Y}\times\{2\}}) = (1-\tau)\mu(\kappa \circ f) + \tau\nu(\kappa \circ g)$. □

COROLLARY 2.2. *Let $\mathscr{E}$, $\tilde{\mathscr{E}}$, $\mathscr{F}$ and $\tilde{\mathscr{F}}$ be dominated experiments such that $\mathscr{E} \sim \tilde{\mathscr{E}}$ and $\mathscr{F} \sim \tilde{\mathscr{F}}$. Then*

$$(1-\tau)\mathscr{E} + \tau\mathscr{F} \sim (1-\tau)\tilde{\mathscr{E}} + \tau\tilde{\mathscr{F}}\,.$$

PROOF. Let F be a finite subset of Θ and let $a \in R^F$. Choose a nonnegative function c on Θ such that it works for all four experiments. Using the notations above for $\mathscr{E}$ and $\mathscr{F}$ and adapting these notations to $\tilde{\mathscr{E}}$ and $\tilde{\mathscr{F}}$ in the obvious way, we find

$$\begin{aligned}||\textstyle\sum_F a_\theta \sigma_\theta|| &= (1-\tau)||\textstyle\sum_F a_\theta \mu_\theta|| + \tau||\textstyle\sum_F a_\theta \nu_\theta|| \\ &= (1-\tau)||\textstyle\sum_F a_\theta \tilde{\mu}_\theta|| + \tau||\textstyle\sum_F a_\theta \tilde{\nu}_\theta|| = ||\textstyle\sum_F a_\theta \tilde{\sigma}||\,.\end{aligned}$$

Hence equivalence follows from Theorem 1.1. □

An experiment $\mathscr{G}$ will be called extremal if $\mathscr{G} \sim \frac{1}{2}\mathscr{E} + \frac{1}{2}\mathscr{F}$ imply $\mathscr{E} \sim \mathscr{F}$. The extremal experiments which are dominated are characterized in

THEOREM 2.3. *A dominated experiment $\mathscr{G}$ is extremal if and only if it admits a boundedly complete and sufficient sub-σ-algebra.*

The proof is a consequence of Propositions 2.4, 2.5 and 2.6 below.

PROPOSITION 2.4. *Let $\tau \in]0, 1[$. Then $(1-\tau)\mathscr{E} + \tau\mathscr{F}$ is dominated if and only if $\mathscr{E}$ and $\mathscr{F}$ are dominated.*

PROOF. We use the same notations as in the definition immediately before Proposition 2.1. Let $c \geqq 0$ be a function on Θ such that $\sigma = \sum c_\theta \sigma_\theta$ is an

L-measure for $(1-\tau)\mathscr{E}+\tau\mathscr{F}$. As in the proof of Proposition 1 we get

$$1=\|\sigma\|=(1-\tau)\textstyle\sum_\theta c_\theta\|\mu_\theta\|+\sum_\theta c_\theta\|\nu_\theta\|\,.$$

It follows that $\mu=\sum_\theta c_\theta\mu_\theta$ and $\nu=\sum_\theta c_\theta\nu_\theta$ are probability measures and that $\sigma([A\times\{1\}]\cup[B\times\{2\}]=(1-\tau)\mu(A)+\tau\nu(B)$. Suppose $\mu(A)=0$. Then $\sigma([A\times\{1\}]+[\varnothing\times\{2\}])=(1-\tau)\mu(A)+\tau\nu(\varnothing)=0$. Hence

$$(1-\tau)\mu_\theta(A)\equiv_\theta\sigma_\theta([A\times\{1\}]+[\varnothing\times\{2\}])\equiv_\theta 0\,.$$

It follows that $\mu\gg\{\mu_\theta:\theta\in\Theta\}$. Similarly $\nu\gg\{\nu_\theta:\theta\in\Theta\}$. This proves the "only if" and the "if" follows from Proposition 2.1. □

PROPOSITION 2.5. *Let $\tau\in\,]0,1[$ and suppose $(1-\tau)\mathscr{E}+\tau\mathscr{F}$ is dominated, and that it admits a boundedly complete and sufficient sub-σ-algebra. Then $\mathscr{E}\sim\mathscr{F}$.*

PROOF. We use the notations of Proposition 2.1. Suppose we have proved the proposition when $\mathscr{E}$ and $\mathscr{F}$ are minimal sufficient. Let $\tilde{\mathscr{E}}$ and $\tilde{\mathscr{F}}$ be minimal sufficient experiments equivalent with, respectively, $\mathscr{E}$ and $\mathscr{F}$. Then $(1-\tau)\tilde{\mathscr{E}}+\tau\tilde{\mathscr{F}}$ is dominated, it is equivalent with $(1-\tau)\mathscr{E}+\tau\mathscr{F}$ by Corollary 2.2, and it admits a boundedly complete and sufficient sub-σ-algebra. By assumption $\tilde{\mathscr{E}}\sim\tilde{\mathscr{F}}$. Hence $\mathscr{E}\sim\mathscr{F}$.

It follows that we may, without loss of generality, assume that $\mathscr{E}$ and $\mathscr{F}$ are minimal sufficient. By Proposition 2.1 $\sigma h^{-1}=(1-\tau)\mu f^{-1}+\tau\nu g^{-1}$. Hence $\mu f^{-1}\ll\sigma h^{-1}$. Let s be a version of $d\mu f^{-1}/d\sigma h^{-1}$. We may specify s so that $0\leqq s\leqq(1-\tau)^{-1}$. It follows that $\mu_\theta(\chi)=\int f_\theta\,d\mu=\int x_\theta(\mu f^{-1})(dx)=\int x_\theta s(x)(\sigma h^{-1})(dx)=\int h_\theta s(h)\,d\sigma=\int s(h)\,d\sigma_\theta$. Hence $\int s(h)\,d\sigma_\theta\equiv_\theta\int 1\,d\sigma_\theta$; i.e., $\int[s(h)-1]\,d\sigma_\theta\equiv_\theta 0$.

The σ-algebra generated by h is, by assumption, boundedly complete and $s\circ h-1$ is bounded. It follows that $s(h)=1$ a.e. σ, i.e., $s=1$ a.e. σh^{-1}. Hence $\mu f^{-1}=\sigma h^{-1}=(1-\tau)\mu f^{-1}+\tau\nu g^{-1}$. It follows that $\mu f^{-1}=\nu g^{-1}$ so that $\mathscr{E}\sim\mathscr{F}$. □

PROPOSITION 2.6. *Let $\mathscr{H}$ be a dominated experiment which does not admit a boundedly complete and sufficient sub-σ-algebra and let $\tau\in\,]0,1[$. Then there are nonequivalent experiments $\mathscr{E}$ and $\mathscr{F}$ so that $\mathscr{H}\sim(1-\tau)\mathscr{E}+\tau\mathscr{F}$.*

PROOF. We may, without loss of generality, assume that $\mathscr{H}$ is minimum sufficient. Write $\mathscr{H}=((U,\mathscr{D});\lambda_\theta\in\Theta)$. By assumption there is a bounded measurable function φ on $(U,\mathscr{D})$ so that $\lambda_\theta(\varphi)\equiv_\theta 0$ and $\lambda_\theta(\varphi\neq 0)>0$ for at least one θ, say for $\theta=\theta_0$. Suppose M is a constant so that $-M<\varphi<M$. Let $\xi,\eta\in\,]0,1]$. Then $\xi(\varphi+M)/2M+(1-\xi)$ and $\eta(\varphi+M)/2M$ are test functions and the λ_θ integrals are, respectively, $(\frac{1}{2}\xi+1-\xi)$ and $\frac{1}{2}\eta$. We may adjust ξ or η so that one of these numbers is τ and choose δ accordingly. By this construction $\lambda_{\theta_0}(\delta\neq\tau)>0$.

We are now ready to define the experiments $\mathscr{E}=((\chi,\mathscr{A}),\mu_\theta:\theta\in 0)$ and $\mathscr{F}=((\mathscr{Y},\mathscr{B}),\nu_\theta:\theta\in\Theta)$. Put

$$\chi=\mathscr{Y}=U\,,\qquad \mathscr{A}=\mathscr{B}=\mathscr{D}$$

$$\mu_0(A)=\frac{1}{1-\tau}\int_A(1-\delta)\,d\lambda_\theta\,;\quad A\in\mathscr{A}$$

and

$$\nu_\theta(A) = \frac{1}{\tau} \int_A \delta \, d\lambda_\theta \,; \quad A \in \mathscr{A} .$$

Then

$$\mu_\theta(\chi) = \frac{1}{1-\tau} \int (1-\delta) \, d\lambda_\theta = \frac{1}{1-\tau} (\lambda_\theta(U) - \tau\lambda_\theta(U)) = 1$$

and

$$\nu_\theta(\mathscr{Y}) = \frac{1}{\tau} \int \delta \, d\lambda_\theta = \frac{1}{\tau} \tau\lambda_\theta(U) = 1 .$$

It follows that μ_θ and ν_θ are probability measures.

Let c be a nonnegative function on Θ such that $\lambda = \sum c_\theta \lambda_\theta$ is an L-measure for $\mathscr{H}$. It follows that $\mu = \sum c_\theta \mu_\theta$ and $\nu = \sum c_\theta \nu_\theta$ are finite measures. Put $r_\theta = d\lambda_\theta/d\lambda$; $\theta \in \Theta$. We may assume that $\sum \{c_\theta r_\theta : c_\theta > 0\} = 1$. Then $d\mu_\theta/d\lambda = (1-\delta)/(1-\tau)r_\theta$ and $d\nu_\theta/d\lambda = \delta\tau^{-1}r_\theta$ so that $d\mu/d\lambda = (1-\delta)/(1-\tau)$ and $d\nu/d\lambda = \delta/\tau$. It follows easily that $\mu \gg \mu_\theta : \theta \in \Theta$, $\nu \gg \nu_\theta : \theta \in \Theta$, and that $d\mu_\theta/d\mu = d\nu_\theta/d\nu = r_\theta$. Put $r = (r_\theta : \theta \in \Theta)$, and write $(1-\tau)\mathscr{E} + \tau\mathscr{F} = ((\mathscr{X}, \mathscr{C}), \sigma_\theta : \theta \in \Theta)$ as before. By Proposition 2.1, $\sigma = \sum_\theta c_\theta \sigma_\theta$ is a finite measure dominating $\{\sigma_\theta : \theta \in \Theta\}$. Put $h_\theta = d\sigma_\theta/d\sigma$ and $h = (h_\theta : \theta \in \Theta)$. By Proposition 2.1 again: $\sigma h^{-1} = (1-\tau)\mu r^{-1} + \tau\nu r^{-1}$. Let κ be a bounded measurable function on R^Θ. Then

$$\begin{aligned} \int \kappa(r) \, d\lambda &= (1-\tau) \int \frac{1-\delta}{1-\tau} \kappa(r) \, d\lambda + \tau \int \frac{\delta}{\tau} \kappa(r) \, d\lambda \\ &= (1-\tau) \int \kappa(r) \, d\mu + \tau \int \kappa(r) \, d\nu . \end{aligned}$$

It follows that $\lambda r^{-1} = (1-\tau)\mu r^{-1} + \tau\nu r^{-1}$ so that $\sigma h^{-1} = \lambda r^{-1}$. Hence $(1-\tau)\mathscr{E} + \tau\mathscr{F} \sim \mathscr{H}$. It remains to show that $\mathscr{E} \not\sim \mathscr{F}$. Suppose $\mathscr{E} \sim \mathscr{F}$. Then, by Theorem 1.1, $\mu r^{-1} = \nu r^{-1} = \lambda r^{-1}$. Le t be λr^{-1}-integrable on R^Θ. Then $t \circ r$ is λ-integrable and

$$\begin{aligned} \int \left[\frac{1-\delta}{1-\tau} - \frac{\delta}{\tau}\right] t \circ r \, d\lambda &= \int t \circ r \, d\mu - \int (t \circ r) \, d\nu \\ &= \int t \, d\mu r^{-1} - \int t \, d\nu r^{-1} = 0 . \end{aligned} \tag{2.1}$$

By minimal sufficiency there is an r-measurable test function $\tilde{\delta}$ on $(U, \mathscr{D})$ so that $\lambda(\delta \neq \tilde{\delta}) = 0$. Let w be a measurable function on R^Θ so that $\tilde{\delta} = w \circ r$. (2.1) may then be written:

$$\int \left[\frac{1-w}{1-\tau} - \frac{w}{\tau}\right] t \, d\lambda = 0 . \tag{2.2}$$

Here we may, since w is λr^{-1} integrable, insert $t = (1-w)/(1-\tau) - w/\tau$. Hence

$$\int \left(\frac{1-w}{1-\tau} - \frac{w}{\tau}\right)^2 d\lambda = 0 . \tag{2.3}$$

It follows that $(1-w)/(1-\tau) = w/\tau$ a.e. λr^{-1} i.e., $w = \tau$ a.e. λr^{-1}, so that $\tilde{\delta} = \tau$

a.e. λ. Hence $\delta = \tau$ a.e. λ. This, however, is impossible since $\lambda_{\theta_0} \ll \lambda$ and $\lambda_{\theta_0}(\delta \neq \tau) > 0$. It follows that $\mathscr{E} \not\sim \mathscr{F}$. □

PROOF OF THEOREM 2.3. 1°. Suppose $\mathscr{G}$ is extremal and dominated. By Proposition 2.6 $\mathscr{G}$ must admit a boundedly complete and sufficient sub-σ-algebra.

2°. Suppose $\mathscr{G}$ is dominated and admits a boundedly complete and sufficient sub-σ-algebra. Then Proposition 2.5 implies that $\mathscr{G}$ is extremal. □

It was shown in [18] that, in the case of a finite parameter set, any experiment is a mixture of extremal experiments. We will here show how this may be generalized to the separable experiments here defined.

A subset Θ_0 of Θ will be called a *separant* for the experiment $\mathscr{E} = ((\chi, \mathscr{A}), \mu_\theta : \theta \in \Theta)$ if it is countable and $\{\mu_\theta : \theta \in \Theta_0\}$ is dense in $\{\mu_\theta : \theta \in \Theta\}$. A *separable experiment* is an experiment having a separant. An experiment $\mathscr{E}$ is separable if and only if $V(\mathscr{E})$ is separable. Any separable experiment $\mathscr{E} = ((\chi, \mathscr{A}), \mu_\theta : \theta \in \Theta)$ is dominated, and the converse holds provided $\mathscr{A}$ is separable.

It follows directly from Bahadur's construction ([1]) of a minimal sufficient σ-algebra that any experiment having a countable parameter set admits a sufficient σ-algebra which is separable. Now any σ-algebra which is sufficient for $\mathscr{E}_{\Theta_0}$ where Θ_0 is a separant for $\mathscr{E}$ is sufficient for $\mathscr{E}$. Thus, as has been pointed out by Pfanzagl [17], any separable experiment admits a minimal sufficient σ-algebra which is separable. This follows also directly from the characterization of $V(\mathscr{E})$ as the set of finite measures in $L(\mathscr{E})$ whose Radon–Nikodym derivatives w.r.t. a given L-measure may be specified measurable w.r.t. a given minimal sufficient σ-algebra.

Let μ be any σ finite measure dominating $\mathscr{E}$. Pfanzagl proved in [17] that $\mathscr{E}$ is separable if and only if $f_\theta = dP_\theta/d\mu$ may be specified so that $(x, \theta) \rightsquigarrow f_\theta(x)$ is jointly measurable w.r.t. $\mathscr{A} \times \mathscr{U}$ where $\mathscr{U}$ is the class of all subsets of Θ. If the condition is satisfied then, by [17], $f_\theta : \theta \in \Theta$ may be specified so that $(x, \theta) \rightsquigarrow f_\theta(x)$ is jointly measurable w.r.t. $\mathscr{A} \times \mathscr{T}$ where $\mathscr{T}$ is the smallest σ-algebra making all maps $\theta \rightsquigarrow P_\theta(A)$; $A \in \mathscr{A}$ measurable.

A few simple facts on separable experiments are collected in

PROPOSITION 2.7.

(i) *If $\mathscr{E}$ has separant Θ_0 and there is a linear contraction from $B(\mathscr{E})$ to $B(\mathscr{F})$ mapping μ_θ into ν_θ, then Θ_0 is a separant for $\mathscr{F}$. In particular, separability is preserved by equivalence.*

(ii) *If $\mathscr{E}$ has separant Θ_0 and $\mathscr{G} = (\sigma_\theta : \theta \in \Theta_0)$ is an experiment satisfying $\mathscr{E}_{\Theta_0} \sim \mathscr{G}$, then there is one and only one experiment $\mathscr{F} = (\sigma_\theta : \theta \in \Theta)$ such that $\mathscr{E} \sim \mathscr{F}$.*

PROOF. (i) follows from the fact that a contraction is continuous.

(ii) Let $\mathscr{E} = ((\chi, \mathscr{A}), \mu_\theta : \theta \in \Theta)$ have separant Θ_0 and let $\mathscr{G} = ((\mathscr{X}, \mathscr{C}), \sigma_\theta : \theta \in \Theta_0)$ be an experiment with parameter set Θ_0. Suppose $\mathscr{E}_{\Theta_0} \sim \mathscr{G}$. Let $\theta \in \Theta$. By assumption there is a sequence $\{\theta_n\}$ in Θ_0 so that $||\mu_{\theta_n} - \mu_\theta|| \to 0$. Hence, by isometry, $||\sigma_{\theta_n} - \sigma_{\theta_m}|| = ||\mu_{\theta_n} - \mu_{\theta_m}|| \to 0$ as $m, n \to \infty$. It follows that there is a probability measure σ_θ on $\mathscr{C}$ such that $\|\sigma_{\theta_n} - \sigma_\theta\| \to 0$. Using equivalence once more we see that σ_θ does not depend on the choice of the sequence $\{\theta_n\}$ in Θ_0.

Let F be a finite (nonempty) subset of Θ and let $a = (a_\theta : \theta \in F)$ be a point in R^F. To each $\theta \in F$ there is a sequence $\{\eta_{n,\theta}\}$ in Θ_0 so that $\|\mu_{\eta_{n,\theta}} - \mu_\theta\| \to 0$. By construction:

$$\begin{aligned}\|\textstyle\sum_F a_\theta \sigma_\theta\| &= \lim_n \|\textstyle\sum_F a_\theta \sigma_{\eta_{n,\theta}}\| = (\text{by equivalence}) \lim_n \|\textstyle\sum_F a_\theta \mu_{\eta_{n,\theta}}\| \\ &= \|\textstyle\sum_F a_\theta \mu_\theta\| \,.\end{aligned}$$

It follows that $\mathscr{E} \sim \mathscr{F}$. Uniqueness follows directly from the isometry criterion.

A simple consequence is

PROPOSITION 2.8. *Any separable experiment is equivalent to an experiment with* [0, 1] *as sample space.*

PROOF. Suppose $\mathscr{E} = (\mu_\theta : \theta \in \Theta)$ is an experiment with separant Θ_0. Let c be a nonnegative function on Θ_0 so that $\mu = \sum_\theta c_\theta \mu_\theta$ is an L-measure for $\mathscr{E}$. Specify $f_\theta = d\mu_\theta/d\mu \geqq 0$; $\theta \in \Theta$ so that $\sum \{c_\theta f_\theta : \theta \in \Theta_0\} = 1$, and put $f = (f_\theta : \theta \in \Theta_0)$. Construct an experiment $\mathscr{G} = ((\mathscr{Y}, \mathscr{B}), \nu_\theta : \theta \in \Theta_0)$ as follows:

$\mathscr{Y}$ = the set of all points $0 \leqq y \in [0, \infty[^{\Theta_0}$ such that $\sum \{c_\theta y_\theta : \theta \in \Theta_0\} \leqq 1$;

$\mathscr{B}$ = the class of Borel subsets of $\mathscr{Y}$;

$\nu_\theta = \mu_\theta f^{-1}$ when $\theta \in \Theta_0$.

Then, by the isometry criterion, $\mathscr{E}_{\Theta_0} \sim \mathscr{G}$. The proof follows from Proposition 2.7 and the fact that $(\mathscr{Y}, \mathscr{B})$ is Borel isomorph with [0, 1]. □

Let us turn to the task of decomposing separable experiments into separable extremal experiments. The final result will be an experiment of the form

$$((T \times \chi, \mathscr{S} \times \mathscr{A}); Q_\theta : \theta \in \Theta)\,,$$

where

(i) T and χ are Borel subsets of Polish spaces and $\mathscr{S}$ and $\mathscr{A}$ are, respectively, the classes of Borel subsets of T and χ.

(ii) $Q_\theta(S \times A) = \int_S P_{\theta t}(A)\pi(dt)$ where π is a probability measure on $\mathscr{S}$ and $P_{\theta t}$; $\theta \in \Theta$, $t \in T$ is a family of probability measures on $\mathscr{A}$ such that $t \leadsto P_{\theta,t}(A)$ is measurable when $\theta \in \Theta$ and $A \in \mathscr{A}$.

The basic properties of our construction are stated as

THEOREM 2.9. *Any separable experiment* $\mathscr{E}$ *with separant* Θ_0 *is equivalent with an experiment*

$$(T \times \chi, \mathscr{S} \times \mathscr{A}, Q_\theta : \theta \in \Theta)$$

satisfying (i), (ii) *and*:

(iii) *The map* $(t, x) \leadsto x$ *from* $T \times \chi$ *is minimal sufficient.*

(iv) *The experiments* $((\chi, \mathscr{A}), P_{\theta,t} : \theta \in \Theta)$; $t \in T$ *are boundedly complete and*

$$\{P_{\theta t} : \theta \in \Theta_0\} \gg \{P_{\theta t} : \theta \in \Theta\}\,; \quad t \in T\,.$$

REMARK. It follows from (iii) that $\mathscr{E} \sim ((\chi, \mathscr{A}), Q_\theta(T \times \cdot); \theta \in \Theta)$. Suppose $\mathscr{E}$ has a particular property which implies that $Q_\theta(T \times A_0) \equiv_\theta 1$ for a certain A_0 in $\mathscr{A}$. Then (ii) and (iv) imply that there is a set N in $\mathscr{S}$ so that $\pi(N) = 0$ and $P_{\theta t}(T \times A_0) \equiv_\theta 1$ when $t \notin N$.

PROOF OF THE THEOREM. Let $\mathscr{E} = ((\mathscr{Y}, \mathscr{B}), \mu_\theta : \theta \in \Theta)$ be an experiment with separant Θ_0. The summation set for any sum denoted by $\sum$ will, in this proof, be Θ_0. Let c be an everywhere positive function on Θ_0 such that $\sum c = 1$. Then $\mu \equiv \sum c_\theta \mu_\theta$ is a probability measure dominating $\{\mu_\theta : \theta \in \Theta\}$. Specify $f_\theta = d\mu_\theta/d\mu$; $\theta \in \Theta_0$ so that $f_\theta \geqq 0$ when $\theta \in \Theta_0$ and $\sum c_\theta f_\theta = 1$.

Define χ as the set of points x in R^{Θ_0} such that $x_\theta \geqq 0$; $\theta \in \Theta_0$, $\sum c_\theta x_\theta = 1$. Denote by K the set of points x in R^{Θ_0} such that $x_\theta \geqq 0$; $\theta \in \Theta_0$ and $\sum c_\theta x_\theta \leqq 1$. Then K is compact and metrizable for the topology of pointwise convergence on Θ_0. The subset χ is a G_δ subset of K. It follows that χ is a Polish space. Let $\mathscr{A}$ be the class of Borel subsets of χ.

The set $\mathscr{V}$ of all probability measures V on $\mathscr{A}$ such that $\int x_\theta V(dx) = 1$; $\theta \in \Theta_0$ is convex and it is compact for the weak* topology. It follows from Krein Millman's theorem ([8], page 131) that the set T of extreme points of $\mathscr{V}$ is nonempty. It may (see Phelps [16], page 7) be shown that T is a G_δ subset of $\mathscr{V}$. Let $\mathscr{S}$ denote the class of Borel subsets of T.

We have, so far, constructed measurable spaces $(\chi, \mathscr{A})$ and $(T, \mathscr{S})$ satisfying (i). Each $V \in \mathscr{V}$ defines an experiment $\{V_\theta : \theta \in \Theta_0\}$ where—for each $\theta \in \Theta_0$—V_θ is the probability measure on $\mathscr{A}$ whose Radon–Nikodym derivative w.r.t. V is $x \rightsquigarrow x_\theta$. The projection $(t, x) \rightsquigarrow x$ from $T \times \chi$ onto χ will be denoted by X.

The map f from $(\mathscr{Y}, \mathscr{B})$ to $(\chi, \mathscr{A})$ is measurable. Put $P = \mu f^{-1}$. Then $P \in \mathscr{V}$. Define, for each $\theta \in \Theta_0$, P_θ as the probability measure on $\mathscr{A}$ whose Radon–Nikodym derivative w.r.t. P is $x \rightsquigarrow x_\theta$. Then, by the isometry criterion, $\mathscr{E}_{\Theta_0} \sim ((\chi, \mathscr{A}), P_\theta : \theta \in \Theta_0)$. It follows from Proposition 2.7 that there is one and only one way of defining probability measures $P_\theta : \theta \in \Theta - \Theta_0$ so that $\mathscr{E} \sim ((\chi, \mathscr{A}), P_\theta : \theta \in \Theta)$. Let, for each θ, s_θ be a finite nonnegative version of the Radon–Nikodym derivative dP_θ/dP. We may assume that $s_\theta(x) = x_\theta$ when $\theta \in \Theta_0$ and $x \in \chi$.

By a theorem of Choquet ([16], page 19) there is a probability measure π on $\mathscr{S}$ and a family $P_t : t \in T$ of probability measures in $\mathscr{V}$ such that $t \rightsquigarrow P_t(A)$ is measurable for any $A \in \mathscr{A}$ and $P(A) = \int P_t(A)\pi(dt)$; $A \in \mathscr{A}$. Define for each $\theta \in \Theta_0$ and each $t \in T$ the probability measure $P_{\theta t}$ on $\mathscr{A}$ by

$$P_{\theta t}(A) = \int_A s_\theta \, dP_t ; \quad A \in \mathscr{A} .$$

Then $t \rightsquigarrow P_{\theta t}(a)$ is measurable when $\theta \in \Theta_0$. We may therefore define Q_θ when $\theta \in \Theta_0$ by

$$Q_\theta(S \times A) = \int_S P_{\theta t}(A)\pi(dt) , \qquad S \in \mathscr{S}, \quad A \in \mathscr{A} . \tag{2.4}$$

Similarily define the probability measure Q on $\mathscr{S} \times \mathscr{A}$ by

$$Q(S \times A) = \int_S P_t(A)\pi(dt) , \qquad S \in \mathscr{S}, \quad A \in \mathscr{A} . \tag{2.5}$$

By (2.4),

$$\sum c_\theta Q_\theta(S \times A) = \int_S [\sum c_\theta P_{\theta t}(A)]\pi(dt) = \int_S P_t(A)\pi(dt)\,; \qquad S \in \mathscr{S}, \quad A \in \mathscr{A}\,.$$

Hence

$$Q = \sum c_\theta Q_\theta\,.$$

Furthermore

$$\begin{aligned}\int_{S\times A} s_\theta\, dQ &= \int I_S(t) I_A(x) s_\theta(x) Q(d(t, x)) = \int_S P_{\theta t}(A)\pi(dt) \\ &= Q_\theta(S \times A)\,; \qquad \theta \in \Theta_0\,, \quad S \in \mathscr{S}, \quad A \in \mathscr{A}\,.\end{aligned}$$

Hence

$$dQ_\theta/dQ = s_\theta\,; \quad \theta \in \Theta_0\,.$$

It follows, since $s_\theta(x) = x_\theta$ when $\theta \in \Theta_0$ and $x \in \chi$, that X is minimal sufficient in the experiment $((T \times \chi, \mathscr{S} \times \mathscr{A}), Q_\theta : \theta \in \Theta_0)$. Now

$$\begin{aligned}Q_\theta(X \in A) = Q_\theta(T \times A) &= \int P_{\theta t}(A)\pi(dt) = \int [\int I_A(x) x_\theta P_t(dx)]\pi(dt) \\ &= \int I_A(x) x_\theta P(dx) = P_\theta(A) \qquad \text{when} \quad \theta \in \Theta_0 \quad \text{and} \quad A \in \mathscr{A}\,.\end{aligned}$$

Hence $(T \times \chi, \mathscr{S} \times \mathscr{A}, Q_\theta : \theta \in \Theta_0) \sim \mathscr{E}_{\Theta_0}$. By Proposition 2.7 again there is one and only one way of defining probability measures $Q_\theta : \theta \in \Theta - \Theta_0$ on $\mathscr{S} \times \mathscr{A}$ so that $(T \times \chi, \mathscr{S} \times \mathscr{A}, Q_\theta : \theta \in \Theta) \sim \mathscr{E}$. Clearly X is minimal sufficient in this experiment also.

Let $\theta \in \Theta$. Then there is a sequence $\{\theta_n\}$ in Θ_0 so that $\mu_{\theta_n} \to \mu_\theta$. Hence $P_{\theta_n} \to P_\theta$ so that $\int |s_{\theta_n} - s_\theta|\, dQ = \int |s_{\theta_n} - s_\theta|\, dP = ||P_{\theta_n} - P_\theta|| \to 0$. It follows, since $Q_{\theta_n} \to Q_\theta$, that $dQ_\theta/dQ = s_\theta$; $\theta \in \Theta$.

Let $S \in \mathscr{S}$ and $\theta \in \Theta$. Then

$$\int_S [\int s_\theta\, dP_t]\pi(ds) = \int I_S(t) s_\theta(x) Q(d(t, x)) = \int I_S(t) Q_\theta(d(t, x)) = Q_\theta(S \times \chi)\,.$$

If $\theta \in \Theta_0$, then $Q_\theta(S \times \chi) = \int_S \pi(dt) = \pi(S)$. By separability this extends to any $\theta \in \Theta$. Hence

$$\int_S [\int s_\theta\, dP_t]\pi(dt) = \pi(S)\,; \quad S \in \mathscr{S}$$

so that

$$\int s_\theta\, dP_t = 1 \qquad \text{for} \quad \pi \quad \text{almost all} \quad t \quad \text{in} \quad T\,.$$

Put $S_\theta = \{t : \int s_\theta\, dP_t \neq 1\}$. Then $\pi(S_\theta) = 0$ for each $\theta \in \Theta$ and $S_\theta = \emptyset$ when $\theta \in \Theta_0$.

Let $\theta \in \Theta - \Theta_0$ and suppose $t \notin S_\theta$. Then we define $P_{\theta,t}$ as the probability measure on $\mathscr{A}$ whose Radon–Nikodym derivative w.r.t. P_t is s_θ. If $\theta \in \Theta - \Theta_0$ and $t \in S_\theta$ then we put $P_{\theta,t} = P_t$. We have now defined all the probability measures $P_{\theta,t}$; $\theta \in \Theta$, $t \in T$ so that:

$$\begin{aligned}dP_{\theta,t}/dP_t &= s_\theta \qquad \text{when} \quad t \notin S_\theta \\ P_{\theta,t} &= P_t \qquad \text{when} \quad t \in S_\theta\,.\end{aligned}$$

It follows that $P_{\theta,t}(A)$ is measurable in t for each $\theta \in \Theta$ and $A \in \mathscr{A}$ and that $P_{\theta t} \ll P_t$ for all $\theta \in \Theta$ and $t \in T$. Let $S \in \mathscr{S}$, $A \in \mathscr{A}$ and $\theta \in \Theta$. Then

$$Q_\theta(S \times A) = \int I_S(t) I_A(x) s_\theta(x) Q(d(t, x)) = \int_S [\int_A s_\theta\, dP_t]\, d\pi = \int_S P_{\theta t}(A)\, d\pi\,.$$

Let $t \in T$. Suppose $\{P_{\theta t}: \theta \in \Theta\}$ was not extremal. Then, since it is minimal sufficient, it cannot be boundedly complete. Hence its restriction $\mathscr{H}_t \equiv \{P_{\theta t}: \theta \in \Theta_0\}$ to Θ_0 is not boundedly complete either. It follows, since $\mathscr{H}_t$ is minimal sufficient, that $\mathscr{H}_t$ is not extremal. Hence, there are nonequivalent experiments $\mathscr{H}_t'$ and $\mathscr{H}_t''$ on Θ_0 so that $\mathscr{H}_t \sim \frac{1}{2}\mathscr{H}_t' + \frac{1}{2}\mathscr{H}_t''$. It follows that there are distinct probability measures V_t' and V_t'' in $\mathscr{V}$ so that $P_t = \frac{1}{2}V_t' + \frac{1}{2}V_t''$. This, however, is (since P_t is an extreme point in $\mathscr{V}$) a contradiction. It follows that the experiments $\{P_{\theta t}: \theta \in \Theta\}$ are extremal so that by minimal sufficiency they are boundedly complete. □

The last part of the proof of Theorem 2.9 is a consequence of the following extremality condition for dominated experiments.

THEOREM 2.10. *Let $\mathscr{E} = ((\chi, \mathscr{A}), P_\theta: \theta \in \Theta)$ be a dominated experiment and let c be a nonnegative function on Θ so that $\pi \equiv \sum c_\theta P_\theta$ is an L-measure for $\mathscr{E}$. For each θ let f_θ be a nonnegative and finite version of $dP_\theta/d\pi$. Put $f = (f_\theta: \theta \in \Theta)$. Then $\mathscr{E}$ is extremal if and only if πf^{-1} is an extreme point for the convex set of probability measures ρ on $[0, \infty[^\Theta$ (with the product σ-algebra) which are supported by $\{x: \sum c_\theta x_\theta = 1\}$ and satisfies $\int x_\theta \rho(dx) \equiv_\theta 1$.*

PROOF. Let the convex set described above be denoted by $\mathscr{V}$.

1°. Suppose $\mathscr{E}$ is extremal and that $\pi f^{-1} = \frac{1}{2}\rho' + \frac{1}{2}\rho''$ where $\rho', \rho'' \in \mathscr{V}$. Define ρ_θ' and ρ_θ'' by:

$$[d\rho_\theta'/d\rho']_x = [d\rho_\theta''/d\rho'']_x = x_\theta\,; \quad x \in [0, \infty[^\Theta .$$

Then $\mathscr{H}' = \{\rho_\theta': \theta \in \Theta\}$ and $\mathscr{H}'' = \{\rho_\theta''; \theta \in \Theta\}$ are experiments, and using Proposition 2.1 it is easily checked that $\mathscr{E} \sim \frac{1}{2}\mathscr{H}' + \frac{1}{2}\mathscr{H}''$. By extremality, $\mathscr{H}' \sim \mathscr{H}''$ so that $\rho' = \rho''$. It follows that πf^{-1} is extreme.

2°. Suppose πf^{-1} is extreme. Let $\mathscr{H}'$ and $\mathscr{H}''$ be experiments such that $\mathscr{E} \sim \frac{1}{2}\mathscr{H}' + \frac{1}{2}\mathscr{H}''$. Write $\mathscr{H}' = ((\chi', \mathscr{A}'), P_\theta': \theta \in \Theta')$ and $\mathscr{H}'' = ((\chi'', \mathscr{A}''), P_\theta'': \theta \in \Theta'')$. It is easily seen that $\pi' = \sum c_\theta P_\theta'$ and $\pi'' = \sum c_\theta P_\theta''$ dominates, respectively, $\{P_\theta': \theta \in \Theta\}$ and $\{P_\theta'': \theta \in \Theta\}$. Let f_θ' and f_θ'' be finite nonnegative versions of, respectively, $dP_\theta'/d\pi'$ and $dP_\theta''/d\pi''$. Put $f' = (f_\theta'; \theta \in \Theta)$ and $f'' = (f_\theta''; \theta \in \Theta)$. It follows from Proposition 2.1 that $\pi f^{-1} = \frac{1}{2}\pi f'^{-1} + \frac{1}{2}\pi f''^{-1}$ and, since $\pi f'^{-1}, \pi f''^{-1} \in \mathscr{V}$, $\pi f'^{-1} = \pi f''^{-1}$ so that $\mathscr{H}' \sim \mathscr{H}''$. Hence, by the definition of extremality, $\mathscr{E}$ is extremal. □

3. Completeness of product experiments. We will throughout this section assume that all experiments under consideration are dominated.

It is well known that several important results in the theory of mathematical statistics involve some notion of completeness. The most important ones are bounded completeness, quadratic completeness (to be defined below) and completeness. As an example we mention only the fact that a dominated experiment admits a sufficient and quadratically complete sub-σ-algebra, if and only if any real function on Θ which is unbiasedly estimable within the class of everywhere quadratically integrable variables, has a UMVU estimator. (The "if" was proved by Lehmann and Scheffe in [13]. Most of the recent comprehensive books on mathematical statistics do not mention the "only if," although this important result was proved by Bahadur [2] in 1957.)

Let $p \in [1, \infty]$. An experiment $\mathscr{E} = ((\chi, \mathscr{A}), P_\theta : \theta \in \Theta)$ will be called *p-complete* if any random variable δ in $\bigcap_\theta L_p(\chi, \mathscr{A}, P_\theta)$ such that $P_\theta(\delta) \equiv_\theta 0$ is $= 0$ a.e. P_θ for all $\theta \in \Theta$. A sub-σ-algebra $\mathscr{B}$ of $\mathscr{A}$ will be called *p-complete* if the restriction of $\mathscr{E}$ to $\mathscr{B}$ is p-complete. We may write "complete" instead of "1-complete" and "quadratically complete" instead of "2-complete." Clearly p-completeness implies q-completeness for $q > p$ and p-completeness for any p implies bounded completeness.

If Θ or χ is finite then these notions are all equivalent. If Θ is infinite, however, then as is well known, they all differ and bounded completeness does not, in general, imply ∞-completeness. We include, for the sake of completeness, two examples.

EXAMPLE 3.1. Put $\chi = \{0, 1, 2, \cdots\}$, $\mathscr{A} =$ the class of all subsets of χ and define probability measures P_n; $n = 1, 2, \cdots$ by

$$P_n(0) = 1 - \frac{1}{n}, \qquad P_n(n) = \frac{1}{n}; \quad n = 1, 2, \cdots.$$

Then $\{P_n;\ n = 1, 2, \cdots\}$ is boundedly complete. It is not, however, ∞-complete.

EXAMPLE 3.2. Let $(\chi, \mathscr{A})$ be as in the previous example. Let $p_n \in]0, 1[$, $n = 1, 2, \cdots$ and define for each $n = 1, 2, \cdots$ the probability measure P_n by

$$\begin{aligned} P_n(0) &= 1 - p_n \\ P_n(i) &= 0 \qquad \text{when} \quad 0 < i < n \\ P_n(n+i) &= \frac{1}{2^{i+1}} p_n; \quad i = 0, 1, 2, \cdots. \end{aligned}$$

A random variable δ is here everywhere integrable if and only if it is P_1-integrable, and this is the case if and only if $\sum_{i=0}^\infty 2^{-i}|\delta(i)| < \infty$. δ_i; $i = 0, 1, 2, \cdots$ is an unbiased estimator of zero if and only if $\delta_i = (1 + p_{i+1}^{-1} - 2p_i^{-1})\delta_0$; $i = 1, 2, \cdots$ and $\sum_i |p_{i+1}^{-1} - 2p_i^{-1}|\,|\delta_0| 2^{-i} < \infty$ and $|\delta_0| 2^{-i} p_i^{-1} \to 0$.

Let $\tau \in]1, \infty]$ and put $p_n = 2^{-n/\tau}$; $n = 1, 2, \cdots$. This model is τ-complete since $\sum_{i=1}^\infty |p_{i+1}^{-1} - 2p_i^{-1}|^\tau 2^{-i} = \infty$. It is not, however, since $\sum_{i=1}^\infty |p_{i+1}^{-1} - 2p_i^{-1}|^{\tau'} 2^{-i} = \sum [2^{\tau'/\tau - 1}]^i < \infty$ when $\tau' \in [1, \tau[$, τ'-complete for any τ' in $[1, \tau[$.

p-completeness is not, of course, preserved by equivalence. The property of having a sufficient and p-complete sub-σ-algebra is, however, preserved by equivalence.

PROPOSITION 3.3. *Let $\mathscr{E}$ be an experiment admitting a p-complete and sufficient sub-σ-algebra. Suppose $\mathscr{E} \sim \mathscr{F}$. Then $\mathscr{F}$ admits a p-complete and sufficient sub-σ-algebra.*

REMARK. The proposition and its proof carry over, with obvious modifications, to the situation where any given nonnegative measurable function on R takes the role of the function: $x \rightsquigarrow |x|^p$.

PROOF. Referring to Section 1 for the case $p = \infty$ we restrict attention to $p \in [0, \infty[$.

Write $\mathscr{E} = ((\chi, \mathscr{A}), P_\theta : \theta \in \Theta)$ and $\mathscr{F} = ((\mathscr{Y}, \mathscr{B}), Q_\theta : \theta \in \Theta)$. Let c be a nonnegative function on Θ so that $\pi \equiv \sum c_\theta P_\theta$ and $\rho \equiv \sum c_\theta Q_\theta$ are L-measures for, respectively, $\mathscr{E}$ and $\mathscr{F}$. Write

$$f_\theta = dP_\theta/d\pi\,; \quad \theta \in \Theta\,, \qquad g_\theta = dQ_\theta/d\rho\,; \quad \theta \in \Theta\,,$$
$$f = (f_\theta : \theta \in \Theta) \qquad \text{and} \qquad g = (g_\theta : \theta \in \Theta)\,.$$

By Theorem 1.1: $\pi f^{-1} = \rho g^{-1}$.

Suppose $\varphi \in \bigcap_\theta L_p(Q_\theta)$, that $Q_\theta(\varphi) \equiv 0$ and that φ is of the form $\varphi = \kappa \circ g$ where κ is a real-valued, measurable function on R^Θ. Then

$$\int |\kappa \circ f|^p \, dP_\theta = \int |\kappa \circ f|^p f_\theta \, d\pi = \int x_\theta |\kappa(x)|^p \, d\pi f^{-1} = \int x_\theta |\kappa(x)|^p \, d\rho\, g^{-1}$$
$$= \int g_\theta |\kappa(g)|^p \, d\rho = \int |\kappa(g)|^p \, dQ_\theta < \infty\,.$$

It follows that $\kappa \circ f \in \bigcap_\theta L_p(P_\theta)$ and $\int \kappa \circ f \, dP_\theta \equiv \int \kappa \circ g \, dQ_\theta \equiv 0$. Hence, since the σ-algebra induced by f is p-complete, $\kappa \circ f = 0$ a.e. π so that $\kappa = 0$ a.e. $\pi f^{-1} = \rho g^{-1}$. Consequently $\varphi = \kappa \circ g = 0$ a.e. ρ. $\square$

Suppose we have a dominated model for independent random variables X_1, $X_2, \cdots, X_n$ which admits a p-complete (boundedly complete) and sufficient statistic. Does the model for $X_1, \cdots, X_m$ where $m < n$ admit a p-complete (boundedly complete) and sufficient statistic? This is the main problem of this section. The answer cannot be an unconditional yes since the hypothesis is satisfied whenever one of the n observations is totally informative. It will, however, be shown that the answer is yes provided we impose a mild regularity condition. This condition is automatically satisfied when our observations are identically distributed.

Let $\mathscr{E} = (P_\theta : \theta \in \Theta)$ and $\mathscr{F} = (Q_\theta : \theta \in \Theta)$ be two experiments. We shall say that $\mathscr{F}$ is *regular w.r.t.* $\mathscr{E}$ if $\bigwedge_F P_\theta = 0$ for any finite subset F of Θ such that $\bigwedge_F Q_\theta = 0$. We shall say that $\mathscr{F}$ is *regular* if $\bigwedge_F Q_\theta \neq 0$ for all nonempty finite subsets of Θ. If $\mathscr{F}$ is regular then $\mathscr{F}$ is regular w.r.t. and experiment $\mathscr{E}$. Clearly any homogeneous experiment is regular.

For any experiment $\mathscr{E} = ((\chi, \mathscr{A}); P_\theta : \theta \in \Theta)$ with finite parameter set Θ and any number $t \in R^\Theta$ such that $t \geqq 0$ and $\sum_\theta t_\theta = 1$, put

$$L_{\mathscr{E}}(t) = \int \prod_\theta [dP_\theta / d \textstyle\sum_\theta P_\theta]^{t_\theta} \, d \textstyle\sum_\theta P_\theta\,.$$

The map $t \rightsquigarrow L_{\mathscr{E}}(t)$ is called the *Laplace transform* (*Hellinger transform*) of $\mathscr{E}$. If $dP_\theta/d\lambda = f_\theta$ for some measure λ then $L_{\mathscr{E}}(t) = \int \prod_\theta f_\theta^{t_\theta} \, d\lambda$. The Laplace transform determines the experiment up to equivalence and we shall use the formulas:

$$L_{\prod_{i=1}^n \mathscr{E}_i} = \prod_{i=1}^n L_{\mathscr{E}_i}$$

and

$$L_{\frac{1}{2}\mathscr{E} + \frac{1}{2}\mathscr{F}} = \tfrac{1}{2} L_{\mathscr{E}} + \tfrac{1}{2} L_{\mathscr{F}}\,.$$

Regularity may now be described in terms of Laplace transforms as follows:

$\mathscr{F}$ is regular w.r.t. $\mathscr{E}$ if and only if $L_{\mathscr{E}_F}(t) = 0$ for any nonempty finite subset of Θ and point $t \in [0, \infty[^F$ satisfying $\sum_F t_\theta = 1$, such that $L_{\mathscr{F}_F}(t) = 0$. $\mathscr{F}$ is regular if and only if the Laplace transform of any experiment $\mathscr{F}_F$ where F is a finite nonempty subset of Θ has no zeros. It follows that regularity is preserved by equivalence and moreover that $\mathscr{F}$ is regular w.r.t. $\mathscr{E}$ if $\tilde{\mathscr{F}}$ is regular w.r.t. $\tilde{\mathscr{E}}$, $\mathscr{F} \sim \tilde{\mathscr{F}}$ and $\mathscr{E} \sim \tilde{\mathscr{E}}$.

The statistical interpretation of regularity is, essentially, that we cannot be sure that our observations will exclude some part of Θ.

We will now show that the regular experiments are precisely the experiments which satisfy the cancellation law for products. Note first that we may, without loss of generality, assume that Θ is finite. One way is easy. Suppose $\mathscr{E} \times \mathscr{G} \sim \mathscr{E} \times \mathscr{H}$ and that $\mathscr{E}$ is regular. Then, since $L_{\mathscr{E}}$ never vanishes, $L_{\mathscr{G}} = L_{\mathscr{H}}$ so that $\mathscr{G} \sim \mathscr{H}$. Suppose, on the other hand, that L has zeroes. Then, by considering the Laplace transforms, $\mathscr{E} \times \mathscr{G} \sim \mathscr{E} \times \mathscr{H}$ whenever $\mathscr{G}_F \sim \mathscr{H}_F$ for all proper subsets F of Θ. (This follows since $L_{\mathscr{E}}(t) = 0$ when $t_\theta > 0$ for all θ.) It remains to show the existence of nonequivalent experiments $\mathscr{G}$ and $\mathscr{H}$ such that $\mathscr{G}_F \sim \mathscr{H}_F$ for all proper subsets F of Θ. We do this as follows:

Put $m = \#\Theta$. We may assume $m \geqq 2$. Let $\varepsilon_i : i = 1, 2, \cdots, 2^{m-1}$ be the vertices of $[0, 1]^\Theta$ having an even number of coordinates $= 1$. Let $\eta_i : i = 1, 2, \cdots, 2^{m-1}$ be the remaining vertices. Take $\{1, 2, \cdots, 2^{m-1}\}$ as sample space for both experiments. Then $\mathscr{G} = \{P_\theta\}$ and $\mathscr{H} = \{Q_\theta\}$ satisfy our requirements provided we put $P_\theta(i) = 2^{2-m}\varepsilon_{\theta i}$ and $Q_\theta(i) = 2^{2-m}\eta_{\theta i}$. Here $\varepsilon_{\theta i}$ and $\eta_{\theta i}$ are the θth coordinate of, respectively, ε_i and η_i.

If we do not want to assume regularity, then cancellation may often be carried out by noting that $\mathscr{E} \times \mathscr{H} \sim \mathscr{E} \times \mathscr{G}$ imply $\mathscr{H} \sim \mathscr{G}$ provided $\mathscr{H}$ and $\mathscr{G}$ are regular w.r.t. $\mathscr{E}$.

Consider first the case of extremal product experiments:

THEOREM 3.4. *Let $\mathscr{E}$ and $\mathscr{F}$ be dominated experiments such that $\mathscr{E} \times \mathscr{F}$ is extremal. Then $\mathscr{E}$ is extremal provided $\mathscr{F}$ is regular* w.r.t. $\mathscr{E}$.

PROOF. Let $\mathscr{G}$ and $\mathscr{H}$ be experiments such that $\mathscr{E} \sim \frac{1}{2}\mathscr{G} + \frac{1}{2}\mathscr{H}$. Then $\mathscr{E} \times \mathscr{F} \sim \frac{1}{2}(\mathscr{G} \times \mathscr{F}) + \frac{1}{2}(\mathscr{H} \times \mathscr{F})$. Hence, since $\mathscr{E} \times \mathscr{F}$ is extremal, $\mathscr{G} \times \mathscr{F} \sim \mathscr{H} \times \mathscr{F}$. Let F be a finite nonempty subset of Θ and t a point in $[0, \infty[^F$ such that $\sum_F t_\theta = 1$. Then $L_{\mathscr{G}_F}(t)L_{\mathscr{F}_F}(t) = L_{\mathscr{H}_F}(t) \circ L_{\mathscr{F}_F}(t)$. It follows that $L_{\mathscr{G}_F}(t) = L_{\mathscr{H}_F}(t)$ when $L_{\mathscr{F}_F}(t) > 0$. Suppose $L_{\mathscr{F}_F}(t) = 0$. Then, since $\mathscr{F}$ has w.r.t. $\mathscr{E}$, $0 = L_{\mathscr{E}_F}(t) = \frac{1}{2}L_{\mathscr{G}_F}(t) + \frac{1}{2}L_{\mathscr{H}_F}(t)$. It follows that $L_{\mathscr{G}_F} = L_{\mathscr{H}_F}$ so that $\mathscr{G}_F \sim \mathscr{H}_F$. Hence $\mathscr{G} \sim \mathscr{H}$ and this proves the extremality of $\mathscr{E}$. □

COROLLARY 3.5. *Let $\mathscr{E}$ be a dominated experiment and suppose $\mathscr{E}^n$ is extremal. Then $\mathscr{E}^m$ is extremal when $1 \leqq m \leqq n$.*

PROOF. Suppose the statement is true for $n = k$ and consider the case $n = k + 1$. Then $\mathscr{E}^n \sim \mathscr{E}^k \times \mathscr{E}$. Clearly $\mathscr{E}$ is regular w.r.t. $\mathscr{E}^k$. Hence $\mathscr{E}^k$ is extremal, and by the induction hypothesis this implies that $\mathscr{E}^m$ is extremal for any $m = 1, 2, \cdots, n$. The corollary follows now by induction on n. □

In order to treat the same problem for p-completeness we need a result which is of some interest in itself.

THEOREM 3.6. *Let* $\mathscr{E} = ((\chi, \mathscr{A}), P_\theta : \theta \in \Theta)$ *and* $\mathscr{F} = ((\mathscr{Y}, \mathscr{B}), Q_\theta : \theta \in \Theta)$ *be dominated experiments such that* $\mathscr{F}$ *is regular* w.r.t. $\mathscr{E}$. *Suppose* $\mathscr{E}$ *is minimal sufficient and let* $\mathscr{S}$ *be a minimal sufficient sub-σ-algebra in the product experiment*

$$\mathscr{E} \times \mathscr{F} = ((\chi \times \mathscr{Y}, \mathscr{A} \times \mathscr{B}); P_\theta \times Q_\theta : \theta \in \Theta) .$$

Let δ *be a real-valued everywhere-integrable variable in* $\mathscr{E}$ *such that*

$$E_{P_\theta \times Q_\theta}(\delta \mid \mathscr{S}) = 0 \qquad \text{a.e.} \quad P_\theta \times Q_\theta ; \quad \theta \in \Theta .$$

Then

$$P_\theta(\delta = 0) \equiv_\theta 1 .$$

PROOF. Choose a countable subset Θ_0 of Θ so that

$$\{P_\theta : \theta \in \Theta_0\} \gg \{P_\theta : \theta \in \Theta\}$$

$$\{Q_\theta : \theta \in \Theta_0\} \gg \{Q_\theta : \theta \in \Theta\}$$

and

$$\{P_\theta \times Q_\theta : \theta \in \Theta_0\} \gg \{P_\theta \times Q_\theta : \theta \in \Theta\} .$$

The summation set for any sum $\sum$ appearing in this proof will—if not otherwise indicated—be Θ_0. Let c be an everywhere positive function Θ_0 such that $\sum c_\theta = 1$ and $\sum c_\theta P_\theta(|\delta|) < \infty$. Put $\pi = \sum c_\theta P_\theta$, $\rho = \sum c_\theta Q_\theta$, $\sigma = \sum c_\theta(P_\theta \times Q_\theta)$, $f_\theta = dP_\theta d\pi$, $f = (f_\theta : \theta \in \Theta)$, $g_\theta = dQ_\theta/d\rho$ and $h_\theta = d(P_\theta \times Q_\theta)/d\sigma$. We may assume that, for each θ, f_θ, g_θ and h_θ are finite nonnegative versions such that h_θ is $\mathscr{S}$-measurable and $\sum_\theta c_\theta f_\theta = 1$. It follows from the minimal sufficiency of $\mathscr{A}$ that we may assume that δ is of the form $\delta = \eta \circ f$ where η is a real-valued measurable function on R^Θ.

By assumption $E_{P_\theta \times Q_\theta} \delta \equiv_\theta 0$ so that $E_{P_\theta} \delta^+ \equiv E_{P_\theta} \delta^-$. We may therefore define for each θ probability measures S_θ and T_θ on $\mathscr{A}$ by:

$$dS_\theta/dP_\theta = a_\theta^{-1}(\delta^+ + 1) \qquad \text{and} \qquad dT_\theta/dP_\theta = a_\theta^{-1}(\delta^- + 1)$$

where $a_\theta = E_{P_\theta}(\delta^+ + 1) = E_{P_\theta}(\delta^- + 1)$. Write $\alpha = \sum c_\theta S_\theta$ and $\beta = \sum c_\theta T_\theta$. Then α and β are probability measures dominating, respectively, $\{S_\theta : \theta \in \Theta\}$ and $\{T_\theta : \theta \in \Theta\}$. Put

$$s_\theta = dS_\theta/d\alpha , \qquad s = (s_\theta : \theta \in \Theta) , \qquad t_\theta = dT_\theta/d\beta \qquad \text{and} \qquad t = (t_\theta : \theta \in \Theta) .$$

Simple calculations yield:

$$f_\theta = a_\theta s_\theta / \textstyle\sum c_\theta a_\theta s_\theta = a_\theta t_\theta / \sum c_\theta a_\theta t_\theta \qquad \text{a.e.} \quad \pi ; \quad \theta \in \Theta .$$

$$(3.1) \qquad \frac{d\pi}{d\alpha} = \left[(\delta^+ + 1) \sum \left(\frac{c_\theta}{a_\theta}\right) f_\theta \right]^{-1} \quad \text{and}$$

$$\frac{d\pi}{d\beta} = \left[(\delta^- + 1) \sum \left(\frac{c_\theta}{a_\theta}\right) f_\theta \right]^{-1} .$$

Define experiments $\mathscr{G}$ and $\mathscr{H}$ by

$$\mathscr{G} = ((\chi, \mathscr{A}), S_\theta : \theta \in \Theta) \qquad \text{and} \qquad \mathscr{H} = ((\chi, \mathscr{A}), T_\theta : \theta \in \Theta) .$$

Suppose $\mathscr{G} \sim \mathscr{H}$. Let b be a bounded measurable function on R^Θ. Then, since $\delta = \eta \circ f$ is π-integrable,

$$\int (b \circ f)[(\eta^+ \circ f) + 1]\, d\pi = \int (b \circ f)(\delta^+ + 1)\, d\pi = \int b \circ f \left[\sum \frac{c_\theta}{a_\theta} f_\theta \right]^{-1} d\alpha$$
$$= \int b(x) \left[\sum \frac{c_\theta}{a_\theta} x_\theta \right]^{-1} (\alpha f^{-1})(dx) ,$$

and similarly

$$\int (b \circ f)[(\eta^- \circ f) + 1]\, d\pi = \int b(x) \left[\sum \frac{c_\theta}{a_\theta} x_\theta \right]^{-1} (\beta f^{-1})(dx) .$$

By Theorem 1.1 $\alpha s^{-1} = \beta t^{-1}$; and by (3.1) $\alpha f^{-1} = \beta f^{-1}$. Hence

$$\int (b \circ f)(\eta \circ f)\, d\pi = \int (b \circ f)[(\eta^+ \circ f) + 1]\, d\pi - \int (b \circ f)[(\eta^- \circ f) + 1]\, d\pi$$
$$= \int b(x) \left[\sum \frac{c_\theta}{a_\theta} x_\theta \right]^{-1} (\alpha f^{-1})(dx)$$
$$- \int b(x) \left[\sum \frac{c_\theta}{a_\theta} x_\theta \right]^{-1} (\beta f^{-1})(dx) = 0 .$$

The validity of this for all bounded measurable functions on R^Θ implies $\delta = \eta \circ f = 0$ a.e. π, i.e., $P_\theta(\delta = 0) \equiv_\theta 1$. The proof will now be completed by showing $\mathscr{G} \sim \mathscr{H}$. We will show this by showing that $L_{\mathscr{G}_F}(t) = L_{\mathscr{H}_F}(t)$ when F is a nonempty finite subset of Θ and $t \in [0, \infty[^F$ satisfies $\sum_F t_\theta = 1$. Consider such a pair (t, F). Then, since δ is σ-integrable,

$$\int \{\textstyle\prod_F h_\theta^{t_\theta}\} |\eta \circ f|\, d\sigma = \int \{\prod_F [h_\theta(\sum_F h_\theta)^{-1}]^{t_\theta}\} |\eta \circ f| (\sum_F h_\theta)\, d\sigma < \infty .$$

It follows that $[\prod_F h_\theta^{t_\theta}](\eta \circ f)$ is σ-integrable and consequently it is $P_\theta \times Q_\theta$-integrable whenever $\theta \in \Theta_0$. Let $\theta \in \Theta_0$. Then, since h is $\mathscr{S}$-measurable,

$$\int [\textstyle\prod_F h_\theta^{t_\theta}](\eta \circ f)\, d(P_\theta \times Q_\theta)$$
$$= \int [\textstyle\prod_F h_\theta^{t_\theta}][E_{P_\theta \times Q_\theta}(\delta \mid \mathscr{S})]\, d(P_\theta \times Q_\theta) = 0 \qquad \text{when} \quad c_\theta > 0 .$$

It follows that $\int [\prod_F h_\theta^{t_\theta}](\eta \circ f)\, d\sigma = 0$, i.e.,

$$\int [\textstyle\prod_F h_\theta^{t_\theta}](\eta^+ \circ f)\, d\sigma = \int [\prod_F h_\theta^{t_\theta}](\eta^- \circ f)\, d\sigma . \tag{3.2}$$

Now $h_\theta = (f_\theta \otimes g_\theta)[\sum c_\theta(f_\theta \otimes g_\theta)]^{-1}$ a.e. σ and $d\sigma \mid d(\pi \times \rho) = \sum c_\theta(f_\theta \otimes g_\theta)$ so that (3.2) may be written

$$\int [\textstyle\prod_F [h_\theta(\sum c_\theta(f_\theta \otimes g_\theta))]^{t_\theta}](\eta^+ \circ f) d(\pi \times \rho) = \text{the same expression in } \eta^-$$

or

$$\int [\textstyle\prod_F (f_\theta \otimes g_\theta)^{t_\theta}](\eta^+ \circ f)\, d(\pi \times \rho) = \text{the same expression in } \eta^- .$$

By Fubini's theorem this may be written

$$\int [\textstyle\prod_F f_\theta^{t_\theta}](\eta^+ \circ f)\, d\pi \circ L_{\mathscr{S}_F}(t) = \int [\prod_F f_\theta^{t_\theta}](\eta^- \circ f)\, d\pi \circ L_{\mathscr{S}_F}(t) .$$

Hence

$$\int [\textstyle\prod_F f_\theta^{t_\theta}]\delta^+\, d\pi = \int [\prod_F f_\theta^{t_\theta}]\delta^-\, d\pi \tag{3.3}$$

when $L_{\mathscr{T}_F}(t) > 0$. By regularity, however, this holds also when $L_{\mathscr{S}_F}(t) = 0$.

It follows that

$$\begin{aligned} L_{\mathscr{G}_F}(t) &= \int \textstyle\prod_F [a_\theta^{-1}(\delta^+ + 1)f_\theta]^{t_\theta}\, d\pi \\ &= \textstyle\prod_F a_\theta^{-t_\theta} \int [\prod_F f_\theta^{t_\theta}]\delta^+\, d\pi + \prod_F a_\theta^{-t_\theta} \int [\prod_F f_\theta^{t_\theta}]\, d\pi \\ &= (\text{the same expression in } \delta^-) = L_{\mathscr{H}_F}(t)\,. \end{aligned}$$ □

We are now ready to prove the analogs of Theorem 3.4 and Corollary 3.5 for p-completeness.

THEOREM 3.7. *Let $\mathscr{E}$ and $\mathscr{F}$ be dominated experiments such that $\mathscr{F}$ is regular* w.r.t. $\mathscr{E}$. *Suppose $\mathscr{E} \times \mathscr{F}$ admits a p-complete and sufficient sub-σ-algebra.*
The $\mathscr{E}$ admits a p-complete and sufficient sub-σ-algebra.

REMARK. By the remark after Proposition 3.3 and Jensen's inequality, the theorem and its proof carry over, with the obvious changes, to the case where the function $x \to |x|^p$ is replaced by any nonnegative convex function on R.

PROOF. We may, without loss of generality, assume that $\mathscr{E}$ is minimal sufficient. Write $\mathscr{E} = ((\chi, \mathscr{A}), P_\theta : \theta \in \Theta)$, $\mathscr{F} = ((\mathscr{Y}, \mathscr{B}), Q_\theta : \theta \in \Theta)$ and let $\mathscr{S}$ be a minimal sufficient sub-σ-algebra in $\mathscr{E} \times \mathscr{F}$. Suppose $\delta \in \bigcap_\theta L_p(P_\theta)$ and that $E_\theta \delta \equiv_\theta 0$. Then $E_{P_\theta \times Q_\theta} \delta \equiv_\theta 0$. By sufficiency there is a $\mathscr{S}$-measurable random variable φ on $\chi \times \mathscr{Y}$ so that

$$E_{P_\theta \times Q_\theta}(\delta \,|\, \mathscr{S}) = \varphi \qquad \text{a.e. } \; P_\theta \times Q_\theta\,; \quad \theta \in \Theta\,.$$

Then $\varphi \in \bigcap_\theta L_p(P_\theta \times Q_\theta)$ and

$$(P_\theta \times Q_\theta)(\varphi) = (P_\theta \times Q_\theta)(\delta) = P_\theta(\delta) = 0\,; \quad \theta \in \Theta\,.$$

Hence $\varphi = 0$ a.e. $P_\theta \times Q_\theta$; $\theta \in \Theta$, so that, by Theorem 3.6, $P_\theta(\delta = 0) \equiv_\theta 1$. □

COROLLARY 3.8. *Let $\mathscr{E}$ be a dominated experiment and suppose $\mathscr{E}^n$ admits a sufficient and p-complete sub-σ-algebra.*
Then $\mathscr{E}^m$: $m = 1, 2, \cdots, n$, admits a sufficient and p-complete sub-σ-algebra.

PROOF. This follows from Theorem 3.7 as Corollary 3.5 followed from Theorem 3.4. □

Let us finally consider a few examples of experiments with finite parameter sets.

EXAMPLE 3.9. Consider a subset Θ of $[0, 1]$ containing m points. Let $\mathscr{E}$ correspond to one binomial trial with unknown success parameter $\theta \in \Theta$. Then $\mathscr{E}^n$ is extremal if and only if $n \leqq m - 1$.

EXAMPLE 3.10. Let X be uniformly distributed on $1, 2, \cdots, \theta$ where $\theta \in \Theta = \{1, 2, \cdots, m\}$. If $\mathscr{E}$ is the experiment defined by X then $\mathscr{E}^n$ is extremal for all n.

EXAMPLE 3.11. Each point $\xi \in [0, 1]^\Theta$ defines an experiment $\mathscr{E}_\xi$ with parameter set Θ where $\mathscr{E}_\xi$ consists in observing a random variable taking the values θ dan Θ with, respectively, probabilities ξ_θ and $1 - \xi_\theta$. Then $L_{\mathscr{E}_\xi}(t) = \prod_\theta \xi_\theta^{t_\theta}$

when $t_\theta < 1$; $\theta \in \Theta$. It follows that

$$\mathscr{E}_\xi \times \mathscr{E}_\eta = \mathscr{E}_{\xi \circ \eta}$$

where $\circ$ indicates pointwise multiplication. In particular,

$$[\mathscr{E}_\xi]^n = \mathscr{E}_{\xi^n} .$$

Now $\mathscr{E}_\xi$ is extremal if and only if $\xi_\theta \equiv_\theta 0$ or $\xi_\theta = 1$ for at least one θ. Hence $\mathscr{E}_\xi{}^n$ is extremal for all n provided $\mathscr{E}_\xi{}^n$ is extremal for some n. Particular cases where $\mathscr{E}_\xi \times \mathscr{E}_\eta$ is extremal but neither $\mathscr{E}_\xi$ nor $\mathscr{E}_\eta$ is extremal may be constructed by choosing points ξ, η in $[0, 1[^\Theta - \{0\}$ such that $\xi \circ \eta = 0$.

REFERENCES

[1] BAHADUR, R. R. (1954). Sufficiency and statistical decision functions. *Ann. Math. Statist.* **25** 432–462.

[2] BAHADUR, R. R. (1957). On unbiased estimates of uniformly minimum variance. *Sankhyā* **18** 211–244.

[3] BIRNBAUM, A. (1960). On the foundations of statistical inference II. IMM-NYU **275**. Institute of Mathematical Sciences, New York Univ.

[4] BIRNBAUM, A. (1961). On the foundation of statistical inference. Binary experiments. *Ann. Math. Statist.* **32** 414–435.

[5] DE GROOT, M. H. (1966). Optimal allocation of observations. *Ann. Inst. Statist. Math.* **18** 13–28.

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

[7] HEYER, H. (1973). *Mathematische Theorie statistischer Experimente.* Springer-Verlag, Berlin.

[8] KELLEY, NAMIOKA (1963). *Linear Topological Spaces.* Van Nostrand, New York.

[9] LE CAM, L. (1964). Sufficiency and approximate sufficiency. *Ann. Math. Statist.* **35** 1419–1455.

[10] LE CAM, L. (1970). Limits of experiments. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 245–262, Univ. of California Press.

[11] LE CAM, L. (1968). Personal communication.

[12] LE CAM, L. (1974). Notes on asymptotic methods in statistical decision theory. Centres de Recherches Math., Univ. de Montréal.

[13] LEHMANN, E. L. and SCHEFFÉ, H. (1950 and 1955). Completeness, similar regions, and unbiased estimation. *Sankhyā* **10** 305–340 and **15** 219–236.

[14] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[15] NEVEU, J. (1965). *Mathematical Foundations of the Calculus of Probability.* Holden-Day, San Francisco.

[16] PHELPS, R. R. (1965). *Lectures on Choquet's Theorem.* Van Nostrand, New York.

[17] PFANZAGL, J. (1969). On the existence of product measurable densities. *Sankhyā Ser. A* **31** 13–18.

[18] TORGERSEN, E. N. (1970). Comparison of experiments when the parameter space is finite. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **16** 219–249.

[19] TORGERSEN, E. N. (1972). Mixtures and completeness properties of dominated pseudo experiments. Statistical research report No. 7, Institute of Mathematics, Univ. of Oslo.

[20] TORGERSEN, E. N. (1975). Notes on comparison of statistical experiments. Statistical memoir No. 1, Institute of Mathematics, Univ. of Oslo.

INSTITUTE OF MATHEMATICS
UNIVERSITY OF OSLO
BLINDERN, OSLO 3
NORWAY