ON THE WEAK COMPACTNESS OF THE SPACE OF EXPERIMENT TYPES

L. Le Cam

University of California, Berkeley

Introduction. Let Θ be a set. An experiment indexed by Θ is often described as a family $\{P_\theta;\ \theta \in \Theta\}$ of probability measures on a given σ-field. Two experiments $\mathcal{E} = \{P_\theta;\ \theta \in \Theta\}$ and $\mathcal{F} = \{Q_\theta;\ \theta \in \Theta\}$ indexed by the same Θ are called equivalent if, for all decision problems, the risk functions achievable for $\mathcal{E}$ or $\mathcal{F}$ are the same.

Following the terminology of [1] we shall call the equivalence class of an experiment $\mathcal{E}$ the type of $\mathcal{E}$. We shall make no notational distinction between experiments and their types. This abuse will not ordinarily lead to excessive confusion.

Let $E(\Theta)$ be the set of all experiment types indexed by a given Θ. A distance, noted Δ was introduced on $\underset{\sim}{E}(\Theta)$ in the paper [2]. The number $\Delta(\mathcal{E},\mathcal{F})$ measures how closely one can match the risk functions available in the two experiments. One can also introduce on $E(\Theta)$ a weak topology as follows.

If $A \subset \Theta$, let $\mathcal{E}_A = P_\theta;\ \theta \in \Theta$. When A is finite, metrize $\underset{\sim}{E}(A)$ by the distance Δ. Define the weak topology of $\underset{\sim}{E}(\Theta)$ as the weakest which renders continuous all the maps $\mathcal{E} \to \mathcal{E}_A$, A finite.

Proposition 2 of [3] asserts that for this weak topology the space $\underset{\sim}{E}(\Theta)$ is a compact Hausdorff space. Unfortunately the proof given there appears to be invalidated by an algebraic mistake, which was pointed out to me by M. W. Moussatat.

If the result itself was incorrect, several arguments in asymptotic theory would have to be reorganized. They would lose in simplicity. Even though the mathematical damage could be considered comparatively slight, the ensuing complications would be a thorough nuisance. What is perhaps even more annoying is that the general theory of comparison of experiments would acquire bizarre features not expressible in terms of finite dimensional distributions of likelihood ratios. These features would appear, in particular and perhaps surprisingly, in the theory of sampling from finite populations.

For these reasons we give here an alternate proof of compactness of $\underset{\sim}{E}(\Theta)$. It is based on the correspondence between experiment types and the conical measures introduced by Choquet (see [4]).

These same conical measures can also be considered as a natural vehicle for the representation of the essential features of distributions of likelihood ratios.

After introduction of appropriate symbols, in Section 2, we give in Section 3 a description of some relations between the conical measures and the distributions of likelihood ratios.

Section 4 uses a transfinite induction argument to construct experiments from canonical measures.

2. Notational definitions. Let $\mathcal{A}$ denote a σ-field carried by a set $\mathcal{X}$. Let A be a finite set. For each $\theta \in A$ let P_θ

be a probability measure on $(\mathcal{X},\mathcal{B})$. Blackwell [5] associates to this family $\mathcal{E} = \{P_\theta;\ \theta \in A\}$ a certain measure β as follows. Take the sum $S = \Sigma_\theta P_\theta$ and select Radon-Nikodym densities dP_θ/ds in such a way that $dP_\theta/dS \geq 0$ and $\Sigma_\theta(dP_\theta/dS) \equiv 1$. Let V be the vector $V = \{(dP_\theta/dS);\ \theta \in A\}$. Consider the finite dimensional space R^A and its unit simplex U formed by vectors $v = \{v_\theta;\ \theta \in A\}$ such that $v_\theta \geq 0$ and $\Sigma v_\theta = 1$.

The likelihood vector V is a measurable map from $(\mathcal{X},\mathcal{B})$ to U. The image of S by V is a finite measure β carried by U. It is called the Blackwell measure of $\mathcal{E} = \{P_\theta;\ \theta \in A\}$. (In fact, Blackwell uses the average $\frac{1}{n}\Sigma_{\theta\in A} P_\theta$, n = card A instead of our sum S. This is an inessential difference.)

The measure β is the sum of the individual distributions $\mathcal{L}(V|\theta)$ of the likelihood vector. It determines these distributions.

Conversely, let β be a positive measure carried by U and such that $\int v_\theta d\beta = 1$ for all $\theta \in A$. To such a Blackwell measure one may associate a family $\{P_\theta;\ \theta \in A\}$, taking for P_θ that measure on R^A which has density v_θ with respect to β. In symbols $dP_\theta = v_\theta d\beta$, or in a simpler notation $P_\theta = v_\theta \cdot \beta$.

Give to R^A its maximum coordinate norm and define a metric on Blackwell measures by the dual Lipschitz norm

$$\|\beta_1 - \beta_2\|_D = \sup_f \left|\int f d(\beta_1 - \beta_2)\right|$$

where f is allowed to range through the set of functions satisfying $|f| \leq 1$ and $|f(v) - f(v')| \leq \|v - v'\|$.

Since U is compact, the set of all Blackwell measures is obviously compact for the metric so defined.

Consider more generally an arbitrary set Θ and an experiment $\mathcal{E} = \{P_\theta;\ \theta \in \Theta\}$ in the sense of [1].

In this text experiments are defined as maps $\theta \to P_\theta$, $P_\theta \geq 0$, $\|P_\theta\| = 1$ from Θ to a space L which is an L-space in the sense of Kakutani [6]. Any experiment in this abstract sense can be represented by a family $\{P_\theta;\ \theta \in \Theta\}$ of probability measures on a suitable space $(\mathcal{X}, \mathcal{B})$. Thus, even though the abstract definition is preferable, we shall proceed here as if $\mathcal{E} = \{P_\theta;\ \theta \in \Theta\}$ was such a family of probability measures.

For infinite sets Θ one cannot readily construct Blackwell measures. The entity which replaces them is a "conical measure" in the sense of Choquet [4]. In the present context it is convenient to introduce these measures as follows.

Consider the space R^Θ of all functions from Θ to the line $R = (-\infty, +\infty)$ and the cone $\Omega = [0,\infty)^\Theta$ which consists of the nonnegative elements of R^Θ. If $\omega \in \Omega$, the value of ω at will be denoted $\omega(\theta)$ and also $\ell_\theta(\omega)$, the symbol ℓ_θ being the name of the linear functional defined on R^Θ by evaluation at θ.

The finite linear combinations $\Sigma_\theta\, c_\theta \ell_\theta$, $\theta \in A$, with A finite form a linear space dual of R^Θ.

We shall call H_0 the space of functions defined on Ω as pointwise suprema of finite sets of linear functions of the type

$\Sigma\, c_\theta \ell_\theta$, $\theta \in A$ with $c_\theta \geq 0$. The linear space H spanned by H_0 is the space of differences $h = h_1 - h_2$ with $h_i \in H_0$.

Definition. A positive linear functional μ defined on H is called a conical measure on Ω. The resultant of μ is the function defined on Θ which takes value $\mu(\ell_\theta)$ at θ.

It will be shown in the following sections that one can establish a one-to-one correspondence between the set $E(\Theta)$ experiment types and the set of conical measures on Ω whose resultant is identically unity.

For the present, note that if Θ is finite and if β is a Blackwell measure on the unit simplex of R^Θ then the integral $\int h d\beta$ is a conical measure of resultant unity on Ω. Conversely any conical measure on Ω has a unique localization on the unit simplex of R^Θ and this localization is a Blackwell measure. Our task, in the infinite case, is to bypass the Blackwell localization to construct directly the bijection between conical measures and experiment types.

Before passing to this note that, for arbitrary Θ, one may metrize the space of conical measures having resultant unity, by a distance between experiments introduced in [2]. The construction is as follows.

By definition an element h of H_0 is a pointwise supremum $h(\omega) = \sup_j h_j(\omega)$ for h_j, $j=1,2,\cdots,k$ finite sums of the form $h_j = \Sigma_\theta\, c_{j,\theta}\ell_\theta$.

We shall say that such an h belongs to H_1 if there is some finite sum $h_0 = \Sigma\, c_{0,\theta}\ell_\theta$ such that (i) $c_{0,\theta} \geq 0$ and $\Sigma_\theta c_{0,\theta} = 1$ and (ii) for all j and ω one has $h_j(\omega) \leq h_0(\omega)$.

If μ_i, $i = 1,2$ are two conical measures of resultant unity one can define their distance by

$$\|\mu_1 - \mu_2\| = \sup\{|\mu_1(h) - \mu_2(h)|;\ h \in H_1\}.$$

It is then easily verifiable that, for finite Θ, this distance is not larger than the dual Lipschitz distance $\|\beta_1 - \beta_2\|_D$ of the corresponding Blackwell measures. Hence the two distances are topologically equivalent (see [7] and [8]). The reasons for the equality between $\|\mu_1 - \mu_2\|$ and the distance of the corresponding experiments are described, in some detail, in [1].

3. Conical measures and likelihood ratios. Take an arbitrary Θ and the space $\Omega = [0, \infty) \subset R^\Theta$ of Section 2. Let $\mathcal{B}$ denote a σ-field carried by some set $\mathcal{X}$. For each $\theta \in \Theta$, let Q_θ be a finite positive measure on $\mathcal{B}$.

If λ is any other positive finite measure on $\mathcal{B}$, one can always define the Radon-Nikodym density $dQ_\theta/d\lambda$ of the part of Q_θ which is dominated by λ.

Consider then some element h of the space H of functions used to define conical measures. The linear functionals which define h all arise from some set $\{\ell_\theta;\ \theta \in A\}$ where A is a finite subset of Θ. Thus h may also be identified to a function on the component $\Omega_A = [0,\infty)^A$ of the product Ω.

Take a λ which dominates all the Q_θ; $\theta \in A$ and compute the integral $\varphi(h) = \int h[v(x)]\lambda(dx)$ where $v(x)$ is the evaluation at $x \in \mathcal{X}$ of the vector $v = \{dQ_\theta/d\lambda;\ \theta \in A\}$ of Radon-Nikodym densities.

Note that every $h \in H$ is positively homogeneous in the sense that for α real and positive one has $h(\alpha\,\omega) = \alpha h(\omega)$.

From this it follows readily that the above integral $\varphi(h)$ remains unchanged if one replaces λ by any other measure λ' which dominates all the Q_θ; $\theta \in A$. Replacing the set A by a larger set A' does not modify the result. Thus, in this manner, we have assigned a number $\varphi(h)$ to each $h \in H$.

<u>Definition</u>. The function φ defined on H by the procedure just described is called the conical measure induced by the family $\{Q_\theta;\ \theta \in \Theta\}$.

We have shown that φ is well defined. It is very easily verified that it is indeed a conical measure and that its resultant is given by the formula

$$\varphi(\ell_\theta) = \|Q_\theta\|.$$

For conical measures defined by such families we shall also need the following easy result.

<u>Lemma</u> 1. <u>Let</u> $Q_{\theta,3} = Q_{\theta,1} + Q_{\theta,2}$ <u>where the</u> $Q_{\theta,i}$ <u>are positive finite measures on</u> $(\mathcal{X},\mathcal{B})$. <u>For each</u> $i=1,2,3$, <u>let</u> φ_i <u>be the corresponding conical measure</u>. <u>Assume that for each</u> $t \in \Theta$ <u>the measure</u> $Q_{t,2}$ <u>is disjoint from all the measures</u> $Q_{\theta,1}$, $\theta \in \Theta$. <u>Then</u> φ_3 <u>is the sum</u> $\varphi_3 = \varphi_1 + \varphi_2$.

<u>Proof</u>. Since each $h \in H$ depends only on a finite set of coordinates ℓ_θ; $\theta \in A$ it is enough to prove this for Θ finite. Let then $\lambda_i = \Sigma_\theta Q_{\theta,i}$ and let V be the vector of densities $V = \{dQ_{\theta,3}/d\lambda_3;\ \theta \in A\}$. Note that the space $\mathcal{X}$ splits into two disjoint parts, say B_1 and B_2 such that on B_i, the density $dQ_{\theta,3}/d\lambda_3$ is equal to $dQ_{\theta,i}/d\lambda_i$. The result follows.

Finally, note that since each $h \in H$ depends only on a finite set $A \subset \Theta$ one can construct the conical measure φ knowing only relations in finite sets $\{Q_\theta;\ \theta \in A\}$. This leads to the following remarks.

Suppose given for each finite $A \subset \Theta$ an experiment type $\mathcal{E}(A) \in E(A)$.

<u>Definition</u>. <u>The family</u> $\{\mathcal{E}(A);\ A \subset \Theta\}$ <u>is called coherent if the inclusion</u> $A_1 \subset A_2$ <u>implies that</u> $\mathcal{E}_{A_1}(A_2)$ <u>of</u> $\mathcal{E}(A_2)$ <u>to</u> A_1.

If $h \in H$ depends only on the coordinates $\theta \in A$, the above described construction assigns to h well-defined number $\varphi(h)$ which depends only on the experiment type $\mathcal{E}(A)$.

Summarizing the relations, one obtains the following lemma.

<u>Lemma</u> 2. <u>The construction described induces a one-to-one correspondence between the set of conical measures of resultant one on</u> $\Omega = [0, \infty)^\Theta$ <u>and the set of coherent families</u> $\{\mathcal{E}(A);\ A \subset \Theta,\ A$ <u>finite</u>$\}$ <u>of experiment types</u>.

4. <u>Constructing an experiment from a conical measure</u>. In this section we assume given a particular set Θ. In order to carry

out some of the arguments it will be convenient to assume that Θ is well ordered and is in fact an initial segment $\Theta = \{\alpha; \alpha < \kappa\}$ of the ordinal numbers. The possibility of such a well ordering depends, of course, on the axiom of choice.

We begin with some lemmas on the structure of the elements h of the space H and on the structure of the conical measures.

Consider a partition $\Theta = \Theta_1 \cup \Theta_2$ of the set Θ and let $C(\Theta_1)$ be the cone formed by elements $\omega \in \Omega = [0, \infty)^\Theta$ such that $\omega(\theta) = 0$ for all $\theta \in \Theta_1$.

Lemma 3. *Let h be an element of H. Then h can be written, in a unique way, as a sum $h = h_1 + h_2$ such that*

(i) h_1 *vanishes on* $C(\Theta_1)$,

(ii) h_2 *is independent of the coordinates* ℓ_θ; $\theta \in \Theta_1$,

(iii) $|h_1|$ *is bounded by a finite sum* $\Sigma\, c_\theta \ell_\theta$; $\theta \in A$, *with* $A \subset \Theta_1$.

Proof. By independence of the ℓ_θ, $\theta \in \Theta_1$ is meant, as usual, that if ω_i, $i = 1,2$ are elements of Ω such that $\ell_\theta(\omega_1) = \ell_\theta(\omega_2)$ for all $\theta \in \Theta_1$; then $h_2(\omega_1) = h_2(\omega_2)$.

Let us first show that the decomposition, if it exists, is necessarily unique. For the purpose, if $\omega \in \Omega$ let ω' be the element of Ω such that $\omega'(\theta) = 0$ for all $\theta \in \Theta_1$ and $\omega'(\theta) = \omega(\theta)$ for all $\theta \in \Theta_2$.

If $h = h_1 + h_2$ is a decomposition satisfying the required conditions, then $h_1(\omega) = 0$ for $\omega \in C(\Theta_1)$ and $h_2(\omega) = h_2(\omega')$. Therefore $h_2(\omega)$ is simply equal to the value $h(\omega')$.

To prove the existence of the decomposition it is enough to consider elements $h \in H$ which are in H_0. Then h has the form

$$h = \sup_j \sum_{\theta \in A} c_{j,\theta} \ell_\theta$$

for some finite set A, some finite set of values $j = 1, 2, \cdots, k$ and for numbers $c_{j,\theta} \geq 0$. Partition A in the form $A = A_1 \cup A_2$ with $A_i = A \cap \theta_i$.

Define $h_{j,i}$ by the formula $h_{j,i} = \Sigma_{\theta \in A_i} c_{j,\theta} \ell_\theta(\omega)$ and let h_2 be defined by

$$h_2(\omega) = \sup_j h_{j,2}(\omega).$$

Then $h(\omega) = h_2(\omega)$ for all $\omega \in C(\theta_1)$. Also, for arbitrary ω, one can write

$$h(\omega) = \sup_j [h_{j,1}(\omega) + h_{j,2}(\omega)]$$

$$\leq \sup_j [h_{j,1}(\omega)] + \sup_j [h_{j,2}(\omega)].$$

Thus $h_1(\omega) = h(\omega) - h_2(\omega) \leq \sup_j [h_{j,1}(\omega)]$, and h_1 is certainly bounded <u>above</u> by a sum $\Sigma_\theta \{c_\theta \ell_\theta;\ \theta \in A_1\}$ with $c_\theta = \Sigma_j c_{j,\theta}$.

On the other hand, since all the coefficients $c_{j,\theta}$ used here are nonnegative, one certainly can write $h(\omega) \geq h_2(\omega)$ for all ω. This gives $h_1 \geq 0$ and the result follows.

<u>Note</u>. The fact that we use only <u>positive</u> linear functionals in the definition of H_0 is used here. One should note however that the space H itself is still a vector lattice for the

pointwise operations. Indeed, if $h \in H$ is written $h = h_1 - h_2$ with $h_i \in H_0$ then $h_1 \vee h_2 \in H_0$ and $h^+ = h \vee 0 = (h_1 \vee h_2) - h_2$ belongs to H.

Going back to the partition $\theta = \theta_1 \cup \theta_2$, consider a conical measure μ on Ω. Following Choquet [4] (Vol. II, page 194) we shall split μ into two parts μ_i, $i=1,2$ with μ_2 "carried" by the cone $C(\theta_1)$ and μ_1 situated "outside" of that cone.

<u>Lemma</u> 4. <u>Let</u> μ <u>be a conical measure on</u> Ω. <u>Then</u> μ <u>can be written</u> (<u>in a unique way</u>) <u>as a sum</u> $\mu = \mu_1 + \mu_2$ <u>of two disjoint conical measures</u> μ_i, $i=1,2$ <u>such that</u> $\mu_2(h) = 0$ <u>for every</u> $h \in H$ <u>which vanishes on</u> $C(\theta_2)$.

<u>Proof</u>. One defines $\mu_1(h)$ for $h \in H$, $h \geq 0$ by the formula

$$\mu_1(h) = \sup_g \{\mu(g) ; 0 \leq g \leq h, \ g(\omega) = 0 \text{ for } \omega \in C(\theta_1)\}.$$

The remainder of the proof consists in verifying that μ_1 and μ_2 have the required properties. This follows from the lattice structure of H.

As explained at the beginning of this section we can assume that θ is a segment $\theta = \{\alpha ; \alpha < \kappa\}$ of the ordinals. Then, for an ordinal $\beta < \kappa$ we can define cones C_β and C_β' as follows:

$$C_\beta = \{\omega ; \ \ell_\theta(\omega) = 0 \text{ for all } \theta < \beta\},$$

$$C_\beta' = \{\omega ; \ \ell_\theta(\omega) = 0 \text{ for all } \theta \leq \beta\}.$$

If μ is a conical measure on Ω, the preceding Lemma 4 allows us to split μ into two disjoint parts, say μ_β and μ_β', with μ_β situated out of C_β and μ_β' carried by C_β.

<u>Lemma</u> 5. <u>Assume that</u> β <u>is a limit ordinal</u>. <u>Then the component</u> μ_β <u>of</u> μ <u>situated out of</u> C_β <u>is the supremum</u> $\mu_\beta = \sup\{\mu_\alpha; \alpha < \beta\}$.

<u>Proof</u>. The inequality $\alpha < \gamma$ implies $C_\alpha \supset C_\gamma$ and therefore $\mu_\alpha \leq \mu_\gamma$. Thus it is clear that $\varphi = \sup\{\mu_\alpha; \alpha < \beta\}$ satisfies the inequality $0 \leq \varphi \leq \mu_\beta$.

Take an $\alpha < \beta$ and a $\theta < \alpha$. Then $\mu_\alpha(\ell_\theta)$ is equal to $\mu(\ell_\theta)$, since ℓ_θ vanishes on C_α. It follows that, for all $\theta < \beta$ the value of the resultant $\varphi(\ell_\theta)$ is precisely equal to $\mu(\ell_\theta)$. Now consider $\psi = \mu_\beta - \varphi$. This is a conical measure whose resultant $\psi(\ell_\theta)$ vanishes for $\theta < \beta$. However, according to Lemma 3 if an element $h \in H$, $h \geq 0$ vanishes on C_β, then $h \leq c\Sigma_\theta\{\ell_\theta; \theta \in A\}$ for some real c and some finite set A contained in the set $\{\alpha; \alpha \times \beta\}$. It follows that $\psi(h) = 0$. The desired result is then a consequence of the definition of μ_β.

To prove the next lemma, we shall need the fact that a conical measure μ on Ω is always σ-smooth. Specifically, if $\{h_n\}$, $n=1,2,\cdots$ is a sequence of elements of H which decreases pointwise to zero on Ω, then $\mu(h_n)$ decreases to zero. This is proved in Choquet [4] (Vol. III, page 19) and results from the fact that the sequence h_n depends only on a countable set of coordinates ℓ_θ. Thus, one can reduce the problem to the case where θ is countable. In this case μ is localizable [4] (Vol. II, page 207) by a Radon measure.

<u>Proposition</u> 1. *Let* μ *be a conical measure on* Ω *and let* β *be a particular ordinal* $\beta \in \Theta$. *Assume that*

(i) μ *is equal to its component carried by* $C_\beta = \{\omega;\ \ell_\theta(\omega) = 0,\ \theta < \beta\}$, *and*

(ii) *the component of* μ *carried by* $C'_\beta = \{\omega;\ \ell_\theta(\omega) = 0;\ \theta \leq \beta\}$ *vanishes.*

Then there is a finite σ-*additive measure* m *carried by the set* $K_\beta = \{\omega:\ \ell_\theta(\omega) = 0$ *for* $\theta < \beta$ *and* $\ell_\beta(\omega) = 1\}$ *such that* $\mu(h) = \int h(\omega)m(d\omega)$ *for all* $h \in H$.

<u>Note</u>. The set $\Omega = [0, \infty)^\Theta$ carries its ordinary product σ-field, say $\mathcal{A}$. The set K_β defined above is not an element of $\mathcal{A}$. The measure m will be defined on the σ-field trace of $\mathcal{A}$ on K_β, or equivalently, on the σ-field defined on Ω by $\mathcal{A}$ and the set K_β itself, with the added requirement that the complement of K_β has measure zero.

<u>Proof</u>. Since μ is carried by C_β two elements h_i, $i=1,2$ of H which agree on C_β yield the same value $\mu(h_1) = \mu(h_2)$. Thus it is sufficient to consider μ on the vector lattice H_β generated by finite sums $\Sigma\{c_\theta\ell_\theta,\ \theta \in A\}$ with $c_\theta \geq 0$ and A contained in the subset $\{\theta;\ \theta \geq \beta\}$ of Θ. This space H_β contains a subspace, say F_β which consists of those $h \in H_\beta$ such that $\ell_\beta(\omega) = 0$ implies $h(\omega) = 0$. In other words, the elements of F_β are the elements of H_β which vanish on C'_β.

Any $f \in F_\beta$ is determined by its restriction f' to K_β according to the relation

$$f(\omega) = \ell_\beta(\omega) f' \frac{\omega}{\ell_\beta(\omega)} \quad .$$

The space F_β' of restrictions of elements of F_β to K_β is a vector lattice which contains the constant function $\ell_\beta'(\omega) \equiv 1$.

Consider a sequence $\{h_n\}$, $h_n \in F_\beta$ which decreases pointwise to zero on K_β. Then h_n decreases to zero pointwise on the entire set Ω.

Define a linear functional m on F_β' by the prescription $m(f') = \mu(f)$. According to the above, this is a well-defined, positive σ-smooth functional. Thus it can be written in the form $m(f') = \int f(\omega) m(d\omega)$ for some σ-additive measure m on K_β. The measure m is well defined on the σ-field of K_β generated by the elements of F_β'. It is easily seen that the σ-field in question is simply the trace on K_β of the product σ-field $\mathcal{A}$.

Now consider an arbitrary element h of H_β. If $h(\omega) \geq 0$ for $\omega \in K_\beta$, then $h(\omega) \geq 0$ for all $\omega \in \Omega$. Indeed $h(\omega) \geq 0$ for all ω which are multiples $\omega = s\omega_1$, $s > 0$, $\omega_1 \in K_\beta$ of elements of K_β, that is, for all $\omega \in \Omega$ such that $\ell_\beta(\omega) > 0$. The positivity of h on Ω itself follows then by continuity. For the same reason, h is well determined by its values on K_β.

Since we have assumed that μ has no component carried by C_β', for every $h \geq 0$, $h \in H_\beta$, one can write

$$\mu(h) = \sup_f\{\mu(f);\ 0 \le f \le h;\ f \in F_\beta\}.$$

In this expression the inequalities $0 \le f \le h$ are supposed to hold on the entire set Ω, but we have just seen that they do so hold already if they are satisfied on K_β itself. Since each $h \in H_\beta$ involves only a finite number of coordinates, the supremum used can be made coutable in each case. It follows that for every $h \in H_\beta$ the value $\mu(h)$ is also equal to the Daniell integral $\mu(h) = \int h(\omega) m(d\omega)$.

This completes the proof of the proposition.

We are now in a position to complete the proof of the bijective correspondence between the set of conical measures and coherent families of experiment types.

For this purpose, let $\mathcal{B}$ denote the σ-field generated on Ω by the product σ-field $\mathcal{A}$ and the sets K_β, $\beta \in \Theta$.

Theorem 1. Let φ be a conical measure on Ω. Assume that φ has resultant $\varphi(\ell_\theta)$ identically equal to unity. Then φ is induced by an experiment $\mathcal{E} = \{P_\theta;\ \theta \in \Theta\}$ formed by probability measures on the σ-field $\mathcal{B}$.

Proof. For each ordinal $\beta \in \Theta$ decompose φ in a sum of disjoint terms $\varphi = \varphi_\beta + \varphi_{\beta,1} + \varphi_{\beta,2}$ where φ_β is the component of φ out of C_β and $\varphi_{\beta,2}$ is the component of φ in C'_β. Then $\varphi_{\beta,1}$ which has no component out of C_β or in C'_β is a candidate for application of Proposition 1. It can be represented through a measure m_β on $K_\beta = \{\omega;\ \ell_\beta(\omega) = 1,\ \ell_\theta(\omega) = 0 \text{ for all } \theta < \beta\}$. by an integral $\varphi_{\beta,1}(h) = \int h(\omega) m_\beta(d\omega)$.

Writing $\ell_\theta \cdot m_\beta$ for the measure which has density ℓ_θ with respect to m_β, consider the sum

$$Q_{\theta,\beta} = \sum_{\alpha<\beta} \ell_\theta \cdot m_\alpha.$$

We claim that for every ordinal β, the component φ_β is induced by the family $\{Q_{\theta,\beta};\ \theta \in \Theta\}$. To show this consider first the case where β is finite, $\beta > 1$ and let S be the sum $S = \Sigma_{\alpha<\beta}\, m_\alpha$. On the σ-field $\mathcal{B}$ this is a sum of disjoint terms. On a set K_α, $\alpha < \beta$ one may write

$$\frac{dQ_{\theta,\beta}}{dS} = \frac{dQ_{\theta,\beta}}{dm_\alpha} = \ell_\theta \ .$$

If $h \in H$ has the form $h = \sup_j \Sigma_\theta\, c_{j,\theta} \ell_\theta$ its restriction to $\cup\{K_\alpha;\ \alpha < \beta\}$ is equivalent, for S, to the expression

$$h = \sup_j \sum_\theta c_{j,\theta} \frac{dQ_{\theta,\beta}}{dS}.$$

It follows that $\int h\, dS = \Sigma_{\alpha<\beta} \int h\, dm_\alpha$ is precisely equal to $\varphi_\beta(h)$.

Proceeding inductively, let γ be a limit ordinal and assume that, for all $\beta < \gamma$, the component φ_β of φ is induced by the sums $Q_{\theta,\beta}$.

Under this assumption, the resultant $\varphi_\beta(\ell_\theta)$ is equal to the norm $\|Q_{\theta,\beta}\|$ for all θ and all $\beta < \gamma$. It follows that $\|Q_{\theta,\gamma}\| = \sup_{\beta<\gamma} \|Q_{\theta,\beta}\| = \sup_{\beta<\gamma} \varphi_\beta(\ell_\theta)$.

According to Lemma 5, this last term is also equal to $\varphi_\gamma(\ell_\theta)$.

Let ν_γ be the conical measure generated by the system $\{Q_{\theta,\gamma}; \theta \in \Theta\}$. The construction implies that if $\beta < \gamma$ then $Q_{\theta,\beta}$ and $Q_{\theta,\gamma} - Q_{\theta,\beta}$ are disjoint. Thus, according to Lemma 1 Section 3, one has $\varphi_\beta \leq \nu_\gamma$ for all $\beta < \gamma$.

Applying Lemma 5 we obtain that

$$\varphi_\gamma = \sup_{\beta<\gamma} \varphi_\beta \leq \nu_\gamma \ .$$

However, since φ_γ and ν_γ have the same resultant $\varphi_\gamma(\ell_\theta) = \nu_\gamma(\ell_\theta) = \|Q_{\theta,\gamma}\|$, the inequality $\varphi_\nu \leq \nu_\gamma$ implies that $\varphi_\gamma = \nu_\gamma$.

Passage from an ordinal β to the next ordinal $\gamma = \beta + 1$ consists in adding to φ_β the term $\varphi_{\beta,1}$ which is represented by the measure m_β. Here again $Q_{\theta,\beta}$ and $\ell_\theta \cdot m_\beta$ are disjoint and the argument given for finite ordinals can be repeated without difficulty.

Thus if φ_β is induced by $\{Q_{\theta,\beta}\}$ for $\beta < \gamma$, then φ_γ is induced by $\{Q_{\theta,\gamma}\}$.

In particular φ itself is induced by the family of probability measures $\{P_\theta; \theta \in \Theta\}$ with

$$P_\theta = \sum_{\alpha\in\Theta} \ell_\theta \cdot m_\alpha \ .$$

This concludes the proof of the theorem.

<u>Corollary</u>. <u>The set of experiment types</u> $\underset{\sim}{E}(\Theta)$ <u>is compact for its weak topology</u>.

Proof. Topologize the space of conical measures on Ω by the topology of pointwise convergence on H. For this topology, the space of conical measures of resultant identically unity is obviously compact. The result follows since the topology in question is very exactly identical to the weak topology of $\underset{\sim}{E}(\Theta)$.

(One could also argue that coherent families $\{\mathcal{E}(A); A \subset \Theta$, A finite$\}$ form a compact space and use the identification between these and the space $\underset{\sim}{E}(\Theta)$. This is essentially the same argument.)

Remark 1. It is clear from the proof of the theorem that the terms $\ell_\theta \cdot m_\beta$ vanish for all $\beta > \theta$. The term $\ell_\theta \cdot m_\beta$ itself represents the part of P_β which is disjoint from all the P_α, $\alpha < \beta$. Each P_θ is, of course, the sum of an at most countable set of terms of the type $\ell_\theta \cdot m_\alpha$.

Remark 2. The above theorem and its corollary refer to the compactness of $\underset{\sim}{E}(\Theta)$ for the weak topology. If Θ is finite this is the same thing as compactness for the experiment distance Δ. However, for infinite Θ the set $\underset{\sim}{E}(\Theta)$ is not compact for Δ. We do not have, at present, very usable criteria for the strong compactness of subsets of $\underset{\sim}{E}(\Theta)$.

Note that a subset S of $\underset{\sim}{E}(\Theta)$ is relatively compact for Δ if and only if its restriction to countable subsets of Θ has the same property. Thus the study of strong compactness could as well be carried out assuming that Θ is countable.

The difference between the weak and strong (that is, Δ) topology can be described on the Choquet measures as follows.

Consider a particular finite sum $h_0 = \Sigma_\theta c_{0,\theta}\ell_\theta$ with $c_{0,\theta} \geq 0$ and $\Sigma c_{0,\theta} = 1$. Let $H(h_0)$ be the subset of H_0 (see Section 2) formed by functions $h \in H_0$ such that $h \leq h_0$. The weak convergence is the topology of uniform convergence on each $H(h_0)$ so constructed. The strong topology is the topology of uniform convergence on the union H_1 of the $H(h_0)$.

Remark 3. The measures P_θ obtained in Theorem 1 are only σ-additive measures. One can easily pass to Radon measures as follows. Let $\overline{\Omega}$ be the compact space $\overline{\Omega} = [0, \infty]^\Theta$ and let $\overline{K}_\beta$ be the set $\overline{K}_\beta = \{\omega; \omega \in \overline{\Omega}, \omega(\beta) = 1, \omega(\theta) = 0 \text{ for } \theta < \beta\}$. Each m_β constructed above has a unique extension $\overline{m}_\beta$ which is a Radon measure on $\overline{K}_\beta$. The introduction of possibly infinite values is largely inessential since for each θ the set $\{\omega; \omega(\theta) = \infty\}$ has measure zero. Thus one could define the sums $\overline{P}_\theta = \Sigma_\alpha \ell_\theta \cdot m_\alpha$ as Radon measures on $\overline{\Omega}$.

Remark 4. The construction carried out in Remark 3 can also be modified to yield a version of the experiment $\mathcal{E} = \{P_\theta; \theta \in \Theta\}$ which is Σ-finite. Explicitly if $\mathcal{E}$ is represented by measures P_θ and a set T with σ-field $\mathcal{A}$, the system is called Σ-finite if there is a partition $\{A_j; j \in J\}$, $A_j \in \mathcal{A}$ of arbitrary cardinality such that (i) when restricted to an A_j the family P_θ is dominated, (ii) any bounded function f whose restriction to each A_j is measurable is already $\mathcal{A}$-measurable and $\int f dP_\theta = \Sigma_j \int_{A_j} f dP_\theta$.

Here this Σ-finiteness can be obtained by taking the compacts $\overline{K}_\beta$ of Remark 3 and letting T be the topological direct sum of the $\overline{K}_\beta$. The space T so obtained is locally compact and the P_θ are still Radon measures. One can take for σ-field $\mathcal{A}$ the σ-field of the universally measurable sets.

This Σ-finiteness is involved in the proof of various representation theorems. One can always obtain it by passing to the Kakutani representation space. However, the space T constructed above may be more directly accessible.

REFERENCES

[1] L. Le Cam. Notes on asymptotic methods in statistical decision theory. Centre de Recherches Mathématiques, Université de Montréal, 1974.

[2] L. Le Cam. "Sufficiency and approximate sufficienty," Ann. Math. Statist., Vol. 35 (1964), pp. 1419-1455.

[3] L. Le Cam. "Limits of experiments," Proceedings Sixth Berkeley Symposium on Math. Stat. and Prob., University of California Press, 1972.

[4] G. Choquet. Lectures on analysis, Benjamin, New York, 1969.

[5] D. Blackwell. "Comparison of experiments," Proc. Second Berkeley Symposium on Math. Stat. and Prob., University of California Press, 1951.

[6] S. Kakutani. "Concrete representation of abstract L-spaces and the mean ergodic theorem," Ann. of Math., Vol. 42 (1941), pp. 523-537.

[7] L. Le Cam. Théorie asymptotique de la décision statistique, Presses de l'Université de Montréal (1969).

[8] E. N. Torgersen. "Comparison of experiments when the parameter space is finite," Zeit f. Wahrscheinlichkeitstheorie u. v. Gebiete, Vol. 16 (1970), pp. 219-249.