Abstract
Let $\mathcal{E}$ be a statistical experiment which is majorized in the sense that there is a non-negative measure such that each probability distribution has a density w.r.t. this measure. If the majorizing measure may be chosen $\sigma$-finite (i.e. $\mathcal{E}$ is dominated), and if any real valued function of the unknown parameter possessing unbiased estimators with finite variance possesses UMVU estimators, then as shown by Bahadur 1957, $\mathcal{E}$ admits a quadratically complete and sufficient $\sigma$-algebra. The purpose of this paper is to show that this extends to any majorized experiment.
1. INTRODUCTION

We shall consider an experiment $\mathcal{E}$ of the traditional measure space type. Thus $\mathcal{E}$ consists of a sample space $(X, \mathcal{A})$ which is just a measurable space, a parameter set $\Theta$ and a rule which to each $\theta \in \Theta$ assigns a probability measure $P_\theta$. Random variables $X$ and $Y$ such that $X = Y$ a.s. $P_\theta$ for all $\theta$ will be identified. Consider also the following two conditions which a given experiment may or may not satisfy:

Condition I Any real valued function of the parameter having an unbiased estimator with everywhere finite variance has a UMVU estimator.

Here UMVU is short for: uniformly minimum variance unbiased.

Condition II The experiment admits a quadratically complete and sufficient $\sigma$-algebra of events.

It follows from the Rao-Blackwell theorem (Rao (1945), Blackwell (1947)) that any experiment satisfying II also satisfies I. If $\mathcal{E}$ is dominated then Bahadur (1957) proved the reverse implication and thus the equivalence of conditions I and II for dominated experiments. We shall here generalize Bahadur's result by showing that conditions I and II are equivalent for majorized experiments.

The following remarks are added for completeness of exposition. Familiarity with the results described below are not needed for the understanding of the proof of the main result.

It was shown by LoCam (1964) that the mathematics of statistical experiments might be greatly simplified by admitting the bounded linear functionals on the band of finite measures gene-
rated by the $P_\theta$'s as random variables. It is of particular interest here that LeCam in this paper was able to demonstrate the existence of a algebra of such functionals which deserved to be called the minimal sufficient algebra.

In order to obtain "generalized" variables which are not necessarily bounded we may proceed by noting that these linear functionals may be represented as uniformly bounded families $(X_\theta : \theta \in \Theta)$ of real random variable satisfying the condition of finite coherence:

To each subset $D$ of $\Theta$ such that $(P_\theta : \theta \in D)$ is dominated corresponds a random variable $X_D$ such that $X_\theta = X_D$ a.s. $P_\theta$ when $\theta \in D$. (Actually it suffices to consider two point sets $D$).

Call a family $(X_\theta : \theta \in \Theta)$ of random variables coherent if there is a random variable $X_\Theta$ such that $X_\theta = X_\Theta$ a.s. $P_\theta$ for all $\theta$. An experiment $\mathcal{E}$ is called coherent if finitely coherent families of real valued random variables are coherent. Some facts on coherence are given in Hasegawa and Perlman (1974), Mussmann (1972), Pitcher (1965), Siebert (1976) and Torgersen (1979). Dominated experiments as well as discrete experiments are coherent. Majorized experiments, however, need not be coherent.

It was shown in Torgersen (1959) that if finitely coherent families of real random variables (not necessarily uniformly bounded) are admitted as "generalized" variables then the natural reformulations of conditions I and II are equivalent without any hypothesis on $\mathcal{E}$ what so ever. If, moreover, the elements in the minimal sufficient algebra of LeCam are all representable as usual random variables, then, as was shown in Torgersen
(1979), this implies that Bahadur's result extends to these cases. The purpose of this paper is to show that this condition may be further relaxed - it suffices to assume that each $P_\theta$ has a minimal support which is representable as a set in $\mathcal{F}$. This, in turn, is equivalent to (as will be explained later) the condition that $\mathcal{E}$ is majorized, i.e. that there is a non negative measure $\mu$ on $\mathcal{F}$ such that each $P_\theta$ has a density w.r.t. $\mu$. 
2. A CONVERSE OF THE RAO-BLACKWELL THEOREM FOR MAJORIZED EXPERIMENTS.

Let the experiment \( \mathcal{E} = (X, \mathcal{A}, P_\theta : \theta \in \Theta) \) be majorized by the non negative measure \( \mu \) on \( \mathcal{A} \), i.e. each \( P_\theta \) has a density \( g_\theta \) w.r.t. \( \mu \). It may, and we shall, without loss of generality be assumed that the densities \( g_\theta \) are all specified finite and non negative. We do not assume that \( \mu \) is \( \sigma \)-finite.

Put for each pair \( (\theta, F) \) such that \( F \) is finite and \( \theta \in F \subseteq \Theta \), \( u_{\theta,F} = g_\theta / \mathbb{E}_{\mathcal{F}}[g_F] \) where \( g_F = \sum_{\theta \in F} g_\theta \). Let \( \mathcal{C} \) be the sub \( \sigma \)-algebra of \( \mathcal{A} \) generated by all functions \( u_{\theta,F} \).

Then \( \mathcal{C} \) is generated by all functions \( u_{\theta,F} \) such that \( \#F \leq 2 \) as well as by the class \( \{ g_{\theta_2}/g_{\theta_1} I_{0, \infty}[g_{\theta_1}] ; \theta_1, \theta_2 \in \Theta \} \).

We shall need the concept of weak closure for sub \( \sigma \)-algebras. Let \( \mathcal{G} \) be any sub \( \sigma \)-algebra of \( \mathcal{A} \). Then the weak closure, \( \mathcal{G}^\wedge \), of \( \mathcal{G} \) is the \( \sigma \)-algebra consisting of the sets \( A \) in \( \mathcal{A} \) having the following property: to each finite sub set \( F \) of \( \Theta \) corresponds a set \( D_F \) in \( \mathcal{G} \) so that \( P_\theta(A \Delta D_F) = 0 \) when \( \theta \in F \). (If \( A \) and \( B \) are sets, then \( A \Delta B = (A-B) \cup (B-A) \) is the symmetric difference of \( A \) and \( B \)).

It is not difficult to see that \( \mathcal{G}^\wedge \) is the closure of \( \mathcal{G} \) within \( \mathcal{A} \) for the topology of pointwise convergence on those finite measures which are dominated by some finite subset of \( \{ P_\theta : \theta \in \Theta \} \). Closure and convergence are, if not stated otherwise, with respect to this topology.

**Theorem** Put \( \mathcal{L} = \mathcal{G}^\wedge \). \( \mathcal{L} \) satisfies condition I if and only if it satisfies condition II. If so, then \( \mathcal{L} \) is sufficient and quadratically complete.
Remark 1. A sub $\sigma$-algebra $\mathcal{G}$ of $\mathcal{A}$ is called quadratically complete if $E_\theta \delta \equiv 0$ and $E_\theta \delta^2 < \infty$ for all $\theta$ imply that $\delta = 0$ a.s. $P_\theta$ for all $\theta$.

Remark 2. It follows from the theorem that $\mathcal{G}$ is minimal sufficient provided $\mathcal{G}$ satisfies I or II. Clearly $\mathcal{G}$ is always pairwise sufficient and it was shown by Siebert (1976) that $\mathcal{G}$ is contained in any other weakly closed and pairwise sufficient sub $\sigma$-algebra.

Remark 3. The implication II $\Rightarrow$ I is a direct consequence of the Rao-Blackwell theorem.

Proof: Let $\mathcal{S}$ be the $\sigma$-algebra generated by the bounded UMVU estimators. By Bahadur (1957), we know that $\mathcal{S}$ is weakly closed and that any $\mathcal{S}$ measurable variable with everywhere finite variance is the UMVU estimator of its expectation. In particular $\mathcal{S}$ is quadratically complete. We shall first show that $\mathcal{S}$ is sufficient and thereby establish the equivalence I $\iff$ II for majorized experiments.

Let $U$ denote the set of real valued random variables with everywhere finite variance and let $T$ ($T_b$) be the class of (bounded) UMVU estimators. By assumption there is a unique expectation preserving map $\Pi$ from $U$ onto $T$. This map is linear and, by Bahadur (1957), $\Pi(\varphi\delta) = \varphi\Pi(\delta)$ whenever $\varphi \in T_b$ and $\delta \in U$.

Suppose we knew that $\Pi$ was non negative: Then $\Pi(\delta) \in T_b$ whenever $\delta$ is bounded. Let $\delta \in U$ and put $\delta_N = \delta[I_{[-N,N]}(\delta)]$. (If $A$ is a set then $I_A$ denotes its indicator function). Then
\( \Pi(\delta_{N}) \rightarrow \Pi(\delta) \) in quadratic mean for \( P_\theta \) for each \( \theta \). Hence, since \( S_3 \) is weakly closed, \( \Pi(\delta) \) is \( S_3 \) measurable. It follows from the results of Bahadur described above that \( \Pi(\delta) \) is a version of \( E_\theta(\delta|S_3) \) for any \( \theta \). Thus \( S_3 \) is sufficient when \( \Pi \) is non-negative. (No assumptions on the existence of a majorizing measure have been used so far.)

If the experiment had been dominated then, by Bahadur (1957), \( \Pi \) would be non-negative. We shall now see how the non-negativity of \( \Pi \) may be deduced by reducing the majorized case to the dominated case.

Consider a function \( u_\theta,F \). Put \( P_F = \sum P_\theta \). Clearly \( u_\theta,F \in U \) and

\[
\int [u_\theta,F - \Pi(u_\theta,F)]^2 dP_F = \int u_\theta,F^2 dP_F + \int (\Pi(u_\theta,F))^2 dP_F - 2 \int u_\theta,F \Pi(u_\theta,F) dP_F
\]

\[
= 2E_\theta[u_\theta,F^2 - \Pi(u_\theta,F)]
\]

Thus \( u_\theta,F = \Pi(u_\theta,F) \) a.s. \( P_F \). Consider any \( \theta' \in \Theta \).

Then \( P_\theta,(|\Pi(u_\theta,F)| < u_\theta,F) = P_\theta,(\mathbb{E}_F > 0 \& |\Pi(u_\theta,F)| < u_\theta,F) = 0 \)

since \( P_\theta,(|\mathbb{E}_F > 0) \cap \mathcal{A}) = 0 \) when \( P_F(\mathcal{A}) = 0 \). Hence \( u_\theta,F^2 \leq \Pi(u_\theta,F)^2 \) so that \( E_\theta'u_\theta,F^2 \leq E_\theta'\Pi(u_\theta,F)^2 \leq E_\theta'u_\theta,F^2 \) for all \( \theta' \in \Theta \). Thus, by the uniqueness of UMVU estimators, \( u_\theta,F = \Pi(u_\theta,F) \) so that each \( u_\theta,F \) is \( S_3 \) measurable. It follows that \( \mathcal{F} \subseteq S_3 \). Fix an element \( \theta_0 \) in \( \Theta \) and put \( B_0 = [\mathbb{E}_{\theta_0} > 0] \).

By the above result, \( u_{\theta_0},[\theta_0] = 1_{B_0} \) is \( S_3 \) measurable so that \( B_0 \in S_3 \). Then the \( P_{\theta_0} \) absolutely continuous part of \( P_\theta \) is the measure: \( \lambda_\theta(A) = P_\theta(AB_0); \ A \in \mathcal{F} \).
Put $\Theta_0 = \{ \theta : \lambda_\theta(x) = P_\theta(B_0) > 0 \}$ and $Q_\theta = \lambda_\theta / \lambda_\theta(x) ; \theta \in \Theta_0$. Consider the experiment $\mathcal{F} = (\chi, \theta : \theta \in \Theta_0)$. $\mathcal{F}$ is clearly dominated by $P_{\Theta_0} = Q_{\Theta_0}$. Let $U_0$, $T_0$ and $T_b$ be defined for $\mathcal{F}$ as $U$, $T$ and $T_b$ was defined for $\mathcal{E}$. We shall now see that $\mathcal{F}$ satisfies condition I.

Clearly $U_0$ consists of those random variables $\delta_0$ such that $\delta I_{B_0} \in U$. Let $\delta \in U_0$. Then $\delta_0 = \Pi(\delta I_{B_0})$ is well defined and $\int \delta_0 d\lambda_\theta = \int I_{B_0} \Pi(\delta I_{B_0}) dP_\theta = \int \Pi(\delta I_{B_0}) dP_\theta = \int \delta I_{B_0} dP_\theta = \int \delta d\lambda_\theta$.

Hence $\int \delta_0 dQ_\theta = \int \delta dQ_\theta ; \theta \in \Theta_0$. Let $Z \in U_0$ be a $\mathcal{F}$-unbiased estimator of zero. Then $Z I_{B_0} \in U$ is an $\mathcal{E}$-unbiased estimator of zero. Hence $\int Z \delta_0 d\lambda_\theta = \int (Z I_{B_0}) \Pi(\delta I_{B_0}) dP_\theta = 0$. Thus $\delta_0$ is $\mathcal{F}$-uncorrelated with any $\mathcal{F}$-unbiased estimator of zero in $U_0$.

It follows that $\delta_0$ is a (and hence the) UMVU estimator of the expectation of $\delta$ in $\mathcal{F}$. Hence $\mathcal{F}$ also satisfies condition I.

By Bahadur (1957), since $\mathcal{F}$ is dominated, the $\mathcal{F}$-UMVU estimator $\delta_0$ of the $\mathcal{F}$-expectation of a non negative $\delta$ in $U_0$ is $\mathcal{F}$-non negative. In particular $I_{B_0} \Pi(\delta) = \Pi(\delta I_{B_0}) \geq 0$ w.r.t. $\mathcal{F}$ when $\delta$ is $\mathcal{E}$-non negative in $U$. Thus, since $P_{\Theta_0} = Q_{\Theta_0}$ and $Q_{\Theta_0}(B_0) = 1$, $\Pi(\delta) \geq 0$; a.s. $P_{\Theta_0}$ when $\delta$ is $\mathcal{E}$-non negative in $U$. Hence, by the arbitrariness of $\theta_0$, $\Pi$ is non negative. Altogether we have shown that $\mathcal{B}$ is sufficient and contains $\mathcal{F}$. It remains to show that $\mathcal{B} \subseteq \mathcal{E}$.

Let $B \in \mathcal{B}$ and fix again an element $\theta_0$ in $\Theta$. Using the notations introduced above we find that the $\mathcal{F}$-UMVU estimator of the $\mathcal{F}$-expectation of $I_{B \cap B_0}$ is $\Pi(I_{B \cap B_0}) = I_{B \cap B_0}$. Hence, by Bahadur (1954, 1957), $B \cap B_0$ belongs to the minimal $\mathcal{F}$-closed, $\mathcal{F}$-sufficient sub $\sigma$-algebra of $\mathcal{A}$.
By the same papers, this $\sigma$-algebra is the $\mathcal{F}$-closure of the $\sigma$-algebra generated by the likelihood ratios $(g_{\theta}/g_{\theta_0})|_{B_0}, \theta \in \Theta_0$. It follows, since these functions are $\mathcal{F}$ measurable, that there is a $S_0$ in $\mathcal{F}$ such that $P_{\theta_0}((B \cap B_0) \triangle S_0) = 0$. We may, since $B_0 \in \mathcal{F}$, assume that $S_0 \subseteq B_0$. Clearly $P_{\theta}((B \cap B_0) \triangle S_0) = 0$. Hence, since $\mathcal{F}$ is closed, $B \cap B_0 \in \mathcal{F}$. Put, for each finite subset $F$ of $\Theta$, $B_F = \bigcup B \cap (g_{\theta} > 0)$. Then $B_F \in \mathcal{F}$. It follows, since $B_F$ converges to $B$ as $F$ increases, that $B \in \mathcal{F}$.

Let us conclude by a few remarks on majorized experiments and minimal supports.

Call a set $A$ in $\mathcal{A}$ a minimal support of the finite measure $A$ if $|\lambda|(A^c) = 0$ and if $P_{\theta}(A-B) = 0$ whenever $|\lambda|(B^c) = 0$. Thus $(g_{\theta} > 0)$ is a minimal support of $P_{\theta}$.

More generally we may say that a variable $\delta$ belonging to some set $K$ of variables is minimal in $K$ if $P_{\theta}(v \geq \delta) = 1$ whenever $v \in K$. Thus $u_{\theta,F}$ is a minimal Radon-Nikodym derivative $dP_{\theta}/dP_F$ where $P_F = \sum_{\theta \in F} P_{\theta}$.

A scrutiny of the arguments used reveals that the essential property needed was the existence of minimal non negative Radon-Nikodym derivatives [We proved that $\Pi(u_{\theta,F}) = u_{\theta,F}$ a.s. $P_F$ so that $|\Pi(u_{\theta,F})|$ is also a version of $dP_{\theta}/dP_F$. Hence, by minimality, $u_{\theta,F} \leq |\Pi(u_{\theta,F})|$ a.s. $P_{\theta}$, for each $\theta', ...]$. This in turn implies and is implied by the existence of minimal supports of each $P_{\theta}$. The latter condition is, however, equivalent to the condition that $\mathcal{G}$ is majorized.

In order to see this let us consider a general, not necessarily majorized experiment $\mathcal{E} = (X, \mathcal{A}; P_{\theta} : \theta \in \Theta)$. Consider the set of finite measures on $\mathcal{A}$ as a vector lattice for the usual setwise
definitions of linear - and lattice operations. [Thus, for example, the greatest lower bound of the two point set \([a, \beta]\) of finite measures is the largest finite measure \(\gamma\) such that \(\gamma \leq a\) and \(\gamma \leq \beta\). This measure will be denoted as \(\gamma = a \wedge \beta\).]

Let \(V\) be the closure of the vector lattice generated by the measures \(P_\theta\) for the total variation norm. It is easily seen that each \(\lambda\) in \(V\) possesses minimal supports provided each \(P_\theta\) possesses minimal supports. By Zorn's lemma there is a maximal disjoint family \((\pi_t : t \in T)\) of probability measures in \(V\).

If \(\theta \in \Theta\) then \(P_\theta\) may be represented as \(P_\theta = P'_\theta + P''_\theta\) where \(P'_\theta\) is dominated by a countable sub set of \(\{\pi_t : t \in T\}\) while \(P''_\theta\) is disjoint from any \(\pi_t\). We may write \(P'_\theta = \sum_t P_{\theta,t}\) where \(P_{\theta,t} = \lim_{n \to \infty} P_\theta \wedge n \pi_t\) is the \(\pi_t\) absolutely continuous part of \(P_\theta\).

Here \(\lim\) is w.r.t. the total variation norm. It follows that \(P'_\theta \in V\). Hence \(P''_\theta = P_\theta - P'_\theta \in V\). If \(P''_\theta \neq 0\) then \(\{\pi_t : t \in T\} \cup \{P''_\theta / P''_\theta(\chi)\}\) is a disjoint set of probability measures in \(V\) which properly contains \(\{\pi_t : t \in T\}\). Hence, by maximality, \(P''_\theta = 0\). It follows that \(P_\theta = \sum_t P_{\theta,t}\); \(\theta \in \Theta\). Put \(\mu = \sum_t \pi_t\) and let \(A_t\) be a minimal support for \(\pi_t\). It is easily seen that \(I_{A_t}\) is a density of \(\pi_t\) w.r.t. \(\mu\). If \(\theta \in \Theta\) then,

\[\Sigma[I_{A_t} dP_\theta / d\pi_t : P_\theta(\chi) > 0],\]

is a density of \(P_\theta\) w.r.t. \(\mu\).

Note that \(\mu\) constructed this way has the additional properties of (See Mussmann (1972)):

**Equivalence:** \(\mu(A) = 0 \iff P_\theta(A) = 0\)

and

**Essentiality:** \(\mu(A) > 0 \Rightarrow 0 < \mu(B) < \infty\)

for some measurable sub set \(B\) of \(A\).
The decomposition above is of independent interest since it provides a method of reducing general experiments to total information experiments and dominated experiments. Thus a sub $\sigma$-algebra $\mathcal{S}$ of $\mathcal{F}$ is sufficient (pairwise sufficient) if and only if it is sufficient (pairwise sufficient) for the experiment $(\pi_t : t \in T)$ and for each of the experiments $\mathcal{E}_t = (P_{\theta t}/P_{\theta t}(x); P_{\theta t}(x) > 0); t \in T$. It is then, of course, essential that the measures $\pi_t$ are all chosen within $V$.

Finally it should be noted that the problem considered here can't be expressed directly in terms of distributions of likelihood ratios since, see LeCam (1964), there is always a $\Sigma$-finite (and hence coherent) experiment yielding the same distributions of likelihood ratios as $\mathcal{E}$ yields.
3 REFERENCES


