COMPLETE CLASS THEOREMS IN EXPERIMENTAL DESIGN

SYLVAIN EHRENFELD
COLUMBIA UNIVERSITY

1. Introduction

There are three broad categories into which problems of experimental design can be classified:

1) the practical problem of deciding which experiments are relevant to the problems under consideration,

2) the analysis of the particular experimental design chosen,

3) the decision as to which of the relevant experiments to perform.

Most of the work in classical design has concerned itself with the first two aspects, while the third has only recently been receiving attention. This paper deals with the third aspect.

Suppose an experimenter has available a family of random variables Y_x depending on a parameter $\theta \in \Omega \subsetneq E^{(p)}$ where $x \in A \subsetneq E^{(k)}$, with A compact and $E^{(p)}$ and $E^{(k)}$ Euclidean spaces. A choice of an experiment of size N is equivalent to choosing N points $x_1, \cdots, x_N$ lying in the set A. Performing the experiment consists in observing $Y_{x_1}, \cdots, Y_{x_N}$. If the experimenter is interested in a set of problems T, concerning the parameter θ, then the question of how to choose $x_1, \cdots, x_N$ becomes important. This is so, since the efficiency and sensitivity of the experiments with regard to the problems in the set T might be very much affected by the choice of $x_1, \cdots, x_N$.

A simple illustration is the following. Suppose Y_{x_α}, $\alpha = 1, \cdots, N$, are uncorrelated random variables with equal variance σ^2, and $E(Y_{x_\alpha}) = \beta_2 + \beta_1 x_\alpha$. The x's are assumed to be fixed constants.

It is known that the variance of the least squares estimate of β_1 is inversely proportional to $\sum (x_\alpha - \bar{x})^2$. Hence, if the values $x_1, \cdots, x_N$ can be chosen in a set $A \subsetneq E^{(1)}$, the experimenter would choose them so that $\sum_\alpha (x_\alpha - \bar{x})^2$ is as large as possible. If one were interested in β_2 as well, it is known that $x_1, \cdots, x_N$ should be chosen so that $\bar{x} = 0$. If A is the interval $-1 \leqq x \leqq 1$, and one were interested in both β_1 and β_2 then, for N even, the observations would be restricted to -1 and $+1$ with half at -1 and the other half at $+1$.

In the above the points $x_1, \cdots, x_N$ were chosen to do "well" in two problems, namely, estimating β_1 and β_2. In general the problems of interest, which we denoted by T, might include estimating certain linear relations of the form $t_1\beta_1 + t_2\beta_2$.

The experimenter can sometimes restrict himself to choosing x's in a subset of A without loss with respect to the problems in the set T. In sections 2 and 3 it will be shown how these subsets can be found.

Sponsored by the Office of Scientific Research of the Air Force, Contract AF 18 (600)-442.

2. The linear hypothesis

The model to be discussed in this section can be stated as follows. Let $Y_{x_1}, \cdots, Y_{x_N}$ be N uncorrelated random variables with common variance σ^2. It is assumed that the expected value of Y_{x_a} is given by

$$E(Y_{x_a}) = \theta_1 x_{a1} + \cdots + \theta_k x_{ak} = \theta' x_a , \qquad a = 1, \cdots, N , \tag{2.1}$$

where $\theta \in \Omega \subsetneq E^{(k)}$ and $x_a = (x_{a1}, \cdots, x_{ak}) \in E^{(k)}$.

The vectors x_a are fixed vectors, and θ is unknown. The coefficients $\theta_1, \cdots, \theta_k$ are the population regression coefficients of $Y' = (Y_{x_1}, \cdots, Y_{x_N})$ on the k vectors $(x_{1J}, \cdots, x_{NJ})$, $J = 1, \cdots, k$, respectively.

In matrix notation the above model can be expressed as

$$E(Y) = x\theta \tag{2.2}$$

where

$$\theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_k \end{pmatrix}, \quad x = \| x_{iJ} \| ; \qquad i = 1, \cdots, N ; \ J = 1, \cdots, k . \tag{2.3}$$

In the experiment of observing $Y_{x_1}, \cdots, Y_{x_N}$ some of the x's might be the same. In order to simplify later proofs, the experiment of observing $Y_{x_1}, \cdots, Y_{x_N}$ will be written as

$$\mathcal{E}_N (n_1, x^{(1)}; \cdots; n_s, x^{(s)}) \qquad \text{with} \sum_{J=1}^{s} n_J = N \tag{2.4}$$

and $x^{(1)}, \cdots, x^{(s)}$ the s different vectors of $x_1, \cdots, x_N$. The experiment $\mathcal{E}_N(n_1, x^{(1)}; \cdots; n_s, x^{(s)})$ is interpreted as an experiment where $Y_{x^{(J)}}$, $J = 1, \cdots, s$, is observed n_J times.

Suppose the experimenter is restricted to choosing the vectors $x^{(J)} \in A \subsetneq E^{(k)}$. A class of experiments $\mathcal{E}_N[A]$ is now defined as follows.

DEFINITION 2.1.

$$\mathcal{E}_N [A] = \{ d \mid d = \mathcal{E}_m (n_1, x^{(1)}; \cdots; n_s, x^{(s)}) \quad \text{with } x^{(J)} \in A \subsetneq E^{(k)} \text{ and } m \leqq N \} . \tag{2.5}$$

The condition $m \leqq N$ is imposed to make the statement of theorems easier.

The class $\mathcal{E}_N[A]$ is the class of experiments where $Y_{x_1}, \cdots, Y_{x_m}$ is observed and $m \leqq N$ and $x_1, \cdots, x_m$ are restricted to the set A.

We now suppose the problem space T to be a set $T \subsetneq E^{(k)}$. A point $(t_1, \cdots, t_k) \in T \subsetneq E^{(k)}$ is interpreted as the problem of estimating $t_1\theta_1 + \cdots + t_k\theta_k$.

Let

$$F(d) = n_1 F_1(d) + \cdots + n_s F_s(d) \tag{2.6}$$

with

$$F_J(d) = \| x_u^{(J)} x_v^{(J)} \| ; \qquad u, v = 1, \cdots, k . \tag{2.7}$$

$F(d)$ is usually called the information matrix associated with experiment d.

Let the variance of the maximum likelihood estimate of $t_1\theta_1 + \cdots + t_k\theta_k$ when experiment d is used be denoted by $V_d[t'\hat{\theta}]$. The maximum likelihood estimate is the same as the least squares estimate in this case, and is the best unbiased linear estimate in the sense of least variance. When t is not estimable the convention of setting $V_d[t'\hat{\theta}] = \infty$ is adopted. We note that t is estimable with respect to d, when $F(d)\rho = t$ has solutions for ρ.

It is known that $V_d[t'\hat{\theta}]$ does not depend on $\theta \in \Omega$ and, when t is estimable with respect to $d = \mathcal{E}_N(n_1, x^{(1)}; \cdots; n_s, x^{(s)})$, that

$$(2.8) \qquad V_d[t'\hat{\theta}] = \sigma^2 \rho_t' F(d) \rho_t$$

where ρ_t is *any* solution of $F(d)\rho_t = t$. When t is not estimable then $\rho_t'F(d)\rho_t$ is set equal to infinity. When $F(d)$ is of full rank, all $t \in E^{(k)}$ are estimable, and $\rho_t'F(d)\rho_t = t'F(d)^{-1}t$.

DEFINITION 2.2. *$\mathcal{E}_L(A_0)$ where $A_0 \subsetneq A$ is said to be essentially complete (T) with respect to $\mathcal{E}_N[A]$ if and only if for any $d \in \mathcal{E}_N[A]$ and any unknown $\theta \in \Omega$ there exists $d^* \in \mathcal{E}_L[A_0]$ such that*

$$(2.9) \qquad V_{d^*}[t'\hat{\theta}] \leqq V_d[t'\hat{\theta}] \qquad \text{for all } t \in T\,.$$

In order to simplify later proofs and to take care of estimability considerations we prove lemma 2.1. The statistical significance of lemma 2.1 is as follows. Let F^* and F be the information matrices associated with experiment d^* and d, respectively. Lemma 2.1 states that if we suppose $t'F^*t - t'Ft \geqq 0$ for all $t \in E^{(k)}$, then

(a) when t is estimable with respect to d, t is estimable with respect to d^*, and

(b) $V_{d^*}[t'\hat{\theta}] \leqq V_d[t'\hat{\theta}]$ for all $t \in E^{(k)}$.

LEMMA 2.1. *If F^* and F are two $k \times k$ nonnegative definite (symmetric) matrices such that*

$$(2.10) \qquad t'F^*t - t'Ft \geqq 0 \qquad \text{for all } t \in E^{(k)}$$

then

(a) *if for any given t, there exists ρ_t such that $F\rho_t = t$, then there exists ρ_t^* such that $F^*\rho_t^* = t$, and*

(b) $\rho_t^{*\prime}F^*\rho_t^* \leqq \rho_t'F\rho_t$.

PROOF.

(a) Let V and V^* be the vector spaces spanned by the column vectors of F and F^*, respectively. Since F^* and F are symmetric matrices, the spaces spanned by the column vectors of F^* and F are the same as the spaces spanned by the row vectors of F^* and F. Then part (a) of lemma 2.1 states that $V^* \supseteq V$. Let us suppose, on the contrary, there exist $l \in V$ and l orthogonal to V^*. Then $F^*l = 0$, and therefore $l'F^*l = 0$. From (2.10) it is seen that $-l'Fl \geqq 0$, and we thus have $l'Fl = 0$, which implies $Fl = 0$. This is a contradiction, since $Fl = 0$ implies l orthogonal to V.

(b) It will first be demonstrated that we can restrict ourselves to F^* and F of the form

$$(2.11) \qquad F^* = \begin{pmatrix} \Lambda & E & 0 \\ E' & I & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad F = \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where

$$(2.12) \qquad \Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix}.$$

This is demonstrated as follows. Let F^*, F be any two matrices satisfying (2.10). It is known there exists a nonsingular matrix D, such that

$$D'F^*D=\begin{pmatrix}A & B\\ B' & C\end{pmatrix}, \qquad D'FD=\begin{pmatrix}I & 0\\ 0 & 0\end{pmatrix}. \tag{2.13}$$

From (2.10) it is seen that A is positive definite, since (2.10) becomes

$$t_1'At_1\geqq t_1't_1 \tag{2.14}$$

where $t'=(t_1, t_2)$. Furthermore, there exist Q, P such that Q,P are orthogonal matrices with

$$Q'AQ=\Lambda, \qquad P'CP=\begin{pmatrix}I & 0\\ 0 & 0\end{pmatrix}. \tag{2.15}$$

We now calculate that

$$\begin{pmatrix}Q & 0\\ 0 & P\end{pmatrix}\begin{pmatrix}A & B\\ B' & C\end{pmatrix}\begin{pmatrix}Q' & 0\\ 0 & P'\end{pmatrix}=\begin{pmatrix}\Lambda & E & G\\ E' & I & 0\\ G & 0 & 0\end{pmatrix}. \tag{2.16}$$

It can be seen that $G=0$, since, for example, the submatrix

$$\begin{pmatrix}\lambda_1 & G_{1n}\\ g_{1n} & 0\end{pmatrix} \tag{2.17}$$

is positive definite, and thus

$$\begin{vmatrix}\lambda & g_{1n}\\ g_{1n} & 0\end{vmatrix}=-g_{1n}^2\geqq 0. \tag{2.18}$$

We have thus demonstrated that F^* and F can be chosen in the indicated form. We now note that

$$\begin{pmatrix}\Lambda & E\\ E' & I\end{pmatrix} \tag{2.19}$$

is positive definite, since from (2.10) we see that

$$(x_1'x_2')\begin{pmatrix}\Lambda & E\\ E' & I\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}\geqq(x_1'x_2')\begin{pmatrix}I & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=x_1'x_1 \tag{2.20}$$

and

$$(0,\ x_2')\begin{pmatrix}\Lambda & E\\ E' & I\end{pmatrix}\begin{pmatrix}0\\ x_2\end{pmatrix}=x_2'x_2. \tag{2.21}$$

We thus derive

$$\rho_t^{*\prime}F^*\rho_t^*=t_1'At_1. \tag{2.22}$$

We know that $A=(\Lambda-E'E)^{-1}$, which yields

$$\rho_t^{*\prime}F^*\rho_t^*=t_1'(\Lambda-E'E)^{-1}t_1. \tag{2.23}$$

Relation (2.10) now becomes

$$t_1'(\Lambda-E'E)^{-1}t_1\leqq t_1't_1 \qquad \text{for all } t_1. \tag{2.24}$$

The last relation will hold if

$$t_1'(\Lambda-E'E)\,t_1-t_1't_1\geqq 0 \qquad \text{for all } t_1 \tag{2.25}$$

This is equivalent to

$$(\Lambda - I - E'E) \tag{2.26}$$

nonnegative definite. From (2.10) we see that

$$\begin{pmatrix} \Lambda & E & 0 \\ E' & I & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \Lambda - I & E & 0 \\ E' & I & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.27}$$

is nonnegative definite, and further that

$$\begin{pmatrix} \Lambda - I & E \\ E' & I \end{pmatrix} \tag{2.28}$$

is positive definite. Because of this R is positive definite, where R is derived from

$$\begin{pmatrix} R & S \\ S' & T \end{pmatrix} = \begin{pmatrix} \Lambda - I & E \\ E' & I \end{pmatrix}^{-1} \tag{2.29}$$

It is known that

$$R = (\Lambda - I - E'E)\,. \tag{2.30}$$

The positive definiteness of R yields the desired results, according to (2.26).

DEFINITION 2.3. *Consider a compact set* $A \subsetneq E^{(k)}$. *If* $x \in A$, *and* $x \neq 0$, *then there exists* $\lambda(x) \geqq 1$ *such that* $\lambda(x)x \in A$, *and if* $\lambda_1 > \lambda(x)$ *then* $\lambda_1 x \notin A$. *With this notation we define* $R(A) \subsetneq A$ *as*

$$R(A) = \{\lambda(x)\,x \mid x \neq 0;\; x \in A\}\,. \tag{2.31}$$

THEOREM 2.1.[1] *If* A *is a compact set in* $E^{(k)}$, *then* $\mathcal{E}_n[R(A)]$ *is essentially complete* (T) *with respect to* $\mathcal{E}_n[A]$ *for all* n, *and all* $T \subsetneq E^{(k)}$.

PROOF. Consider any experiment $d \in \mathcal{E}_n(A)$. Suppose

$$d = \mathcal{E}_n(n_1, x^{(1)}; \cdots; n_s, x^{(s)})\,,\; x^{(J)} \in A\,. \tag{2.32}$$

The information matrix associated with d is $F(d)$ where

$$F(d) = n_1F_1(d) + \cdots + n_sF_s(d) = \sum n_JF_J(d) \tag{2.33}$$

with

$$F_J(d) = \| x_u^{(J)} x_v^{(J)} \|\,. \tag{2.34}$$

Let us consider $\bar{x}^{(J)} = \lambda(x^{(J)})x^{(J)} \in R(A)$ in the definition of $R(A)$. Also let

$$d^* = \mathcal{E}_n(n_1, \bar{x}^{(1)}; \cdots; n_s, \bar{x}^{(s)}) \in \mathcal{E}_n[R(A)]\,. \tag{2.35}$$

Experiment d^* has information matrix

$$F(d^*) = \sum n_JF_J(d^*) \tag{2.36}$$

with

$$F_J(d^*) = \| \lambda^2(x^{(J)})\, x_u^{(J)} x_v^{(J)} \| = \lambda^2(x^{(J)})\,F_J(d)\,. \tag{2.37}$$

[1] This theorem was suggested by Professor T. W. Anderson of Columbia University.

The condition in lemma 2.1 that $t'F(d^*)t - t'F(d)t \geqq 0$ for all $t \in E^{(k)}$ is satisfied since

$$t'F(d^*)t - t'F(d)t = t'[F(d^*) - F(d)]t = \sum_J n_J[\lambda^2(x^{(J)}) - 1]t'F_J(d)t \geqq 0 \tag{2.38}$$

and

$$\lambda(x^{(J)}) \geqq 1, \quad t'F_J(d)t = \left[\sum t_t x_r^{(J)}\right]^2 \geqq 0. \tag{2.39}$$

Now the conclusion of theorem 2.1 follows from lemma 2.1 and the definition of essential completeness.

THEOREM 2.2. *If A is a convex body in $E^{(k)}$ with a total of m extreme points $b^{(1)}, \cdots, b^{(m)}$, then $\mathcal{E}_{n+m}(b^{(1)}, \cdots, b^{(m)})$ is essentially complete (T) with respect to $\mathcal{E}_n[A]$ for all n and all $T \subseteq E^{(k)}$.*

PROOF. Let $d = \mathcal{E}_n(n_1, x^{(1)}; \cdots; n_s, x^{(s)}) \in \mathcal{E}_n(A)$ and let $d^* = \mathcal{E}_r(\bar{n}_1, b^{(1)}; \cdots; \bar{n}_m, b^{(m)})$ with $\sum n_j = r \leqq n + m$ and $\bar{n}_J \neq 0$ for $J = 1, \cdots, m$.

The information matrices associated with experiments d and d^* will be denoted by $F(d)$ and $F(d^*)$, respectively. It will be shown that there exist $\bar{n}_1, \cdots, \bar{n}_m$, such that $\bar{n}_J \neq 0$, and $t'F(d^*)t - t'F(d)t \geqq 0$ for all $t \in E^{(k)}$. This with the use of lemma 2.1 will prove the theorem.

Since $x^{(J)} \in A$ and A is a convex set generated by $b^{(1)}, \cdots, b^{(m)}$, that is,

$$A = \left\{\sum_{i=1}^{m} \lambda_i b^{(i)};\ \lambda_i \geqq 0, i = 1, 2, \cdots, m \quad \text{and} \quad \sum \lambda_i = 1\right\}, \tag{2.40}$$

we have

$$x^{(a)} = \sum_{J=1}^{m} \lambda_J^{(a)} b^{(J)}, \qquad \lambda_J^{(a)} \geqq 0, \sum_J \lambda_J^{(a)} = 1, \quad J = 1, \cdots, m;\ a = 1, \cdots, s. \tag{2.41}$$

Some calculation shows that $F(d)$ and $F(d^*)$ can be written as

$$F(d^*) = \|a_{uv}^*\|, \qquad a_{uv}^* = b_u'Pb_v, \tag{2.42}$$

$$F(d) = \|a_{uv}\|, \qquad a_{uv} = b_u'Mb_v, \tag{2.43}$$

where

$$b_u' = (b_n^{(1)}, \cdots, b_n^{(m)}), \tag{2.44}$$

$$P \equiv \left\|\begin{matrix} \bar{n}_1 & & 0 \\ & \ddots & \\ 0 & & \bar{n}_m \end{matrix}\right\|, \qquad M = \|m_{uv}\|, \tag{2.45}$$

$$m_{uv} = \sum_a n_a \lambda_u^{(a)} \lambda_v^{(a)}. \tag{2.46}$$

It is known from matrix theory that there exists a nonsingular matrix Q such that $b_u = Q\bar{b}_u$ with

$$a_{uv} = b_u'Mb_v = \lambda_1 \bar{b}_u^{(1)} \bar{b}_v^{(1)} + \cdots + \lambda_m \bar{b}_u^{(m)} \bar{b}_v^{(m)} \tag{2.47}$$

$$a_{uv}^* = b_u'Pb_v = \bar{b}_u^{(1)} \bar{b}_v^{(1)} + \cdots + \bar{b}_u^{(m)} \bar{b}_v^{(m)} \tag{2.48}$$

where $\lambda_1, \cdots, \lambda_m$ are the characteristic roots of MP^{-1}. Denoting $X_J = \|\bar{b}_u^{(J)}\bar{b}_v^{(J)}\|$, $J = 1, \cdots, m$, it is seen that $F(d)$ and $F(d^*)$ become

$$F(d) = \lambda_1 X_1 + \cdots + \lambda_m X_m \tag{2.49}$$

$$F(d^*) = X_1 + \cdots + X_m \,. \tag{2.50}$$

It is first noted that X_J is nonnegative since

$$t' X_J t = \sum_{r,\,s} \bar{b}_r^{(J)} \bar{b}_s^{(J)} t_r t_s = \Big[\sum_r t_r \bar{b}_r^{(J)}\Big]^2 \geqq 0 \tag{2.51}$$

and thus

$$t' [F(d^*) - F(d)]\, t = \sum_J (1 - \lambda_J)\, t' X_J t \geqq 0 \tag{2.52}$$

if $\bar{n}_1, \cdots, \bar{n}_m$ can be chosen so that $|\lambda_J| \leqq 1$ for $J = 1, \cdots, m$.

This can be done by choosing $\bar{n}_1, \cdots, \bar{n}_m$ so that the column sums of MP^{-1} are less than or equal to one, since a matrix that has all positive elements, and whose column sums are less than or equal to one, has all characteristic roots $|\lambda_J| \leqq 1$.

Let $\Delta = \|\Delta_{uv}\| = MP^{-1}$, then $\Delta_{uv} \geqq 0$ and

$$\Delta_{uv} = \frac{1}{\bar{n}_v} \sum_a n_a \lambda_u^{(a)} \lambda_v^{(a)} \,. \tag{2.53}$$

The column sums are

$$\sum_u \Delta_{uv} = \frac{1}{\bar{n}_v} \sum_a n_a \lambda_v^{(a)} \,. \tag{2.54}$$

Now $\sum_u \Delta_{uv} \leqq 1$ when $\bar{n}_v = \Big[\sum_a n_a \lambda_v^{(a)}\Big] + 1$, where $[y]$ denotes the greatest integer in y. It is seen that $r = \sum_v \bar{n}_v \leqq n + m$ and $\bar{n}_v \neq 0$. This choice of $\bar{n}_1, \cdots, \bar{n}_m$ therefore satisfies the required conditions.

The following corollaries follow readily.

COROLLARY 2.1. *If A is a compact set in $E^{(k)}$ such that the convex closure of A is generated by m vectors $b^{(1)}, \cdots, b^{(m)} \in A$, then $\mathcal{E}_{n+m}[b^{(1)}, \cdots, b^{(m)}]$ is essentially complete (T) with respect to $\mathcal{E}_n[A]$, for all n, and all $T \subsetneq E^{(k)}$.*

PROOF. The corollary is clear, since $\mathcal{E}_{n+m}[b^{(1)}, \cdots, b^{(m)}]$ is essentially complete with respect to $\mathcal{E}_n[C(A)]$ by theorem 2.2 [$C(A)$ denotes the convex closure of set A], and $\mathcal{E}_n[C(A)]$ is essentially complete with respect to $\mathcal{E}_n(A)$, since $C(A) \supsetneq A$.

COROLLARY 2.2. *Let A be a compact set in $E^{(k)}$ such that $R(A)$ has the property that the convex closure of $R(A)$ is generated by m vectors $b^{(1)}, \cdots, b^{(m)}$ in $R(A)$; then, $\mathcal{E}_{n+m}[b^{(1)}, \cdots, b^{(m)}]$ is essentially complete (T) with respect to $\mathcal{E}_n[A]$ for all n and all $T \subsetneq E^{(k)}$.*

PROOF. The corollary is clear, since $\mathcal{E}_{n+m}[b^{(1)}, \cdots, b^{(m)}]$ is essentially complete (T) with respect to $\mathcal{E}_n\{R(A)\}$ by corollary 2.1, and $\mathcal{E}_n\{R(A)\}$ is essentially complete (T) with respect to $\mathcal{E}_n(A)$ by theorem 2.1.

In an asymptotic sense (as $n \to \infty$), it is true that when A is a convex body in $E^{(k)}$, generated by vectors $b^{(1)}, \cdots, b^{(m)}$, then $\mathcal{E}_n[b^{(1)}, \cdots, b^{(m)}]$ is essentially complete (T) with respect to $\mathcal{E}_n(A)$.

The notion of asymptotic essential completeness will now be defined.

DEFINITION 2.4. *$\mathcal{E}_n[A_0]$ is asymptotically essentially complete (T) with respect to $\mathcal{E}_n[A]$*

if and only if for any sequence of experiments $d_{N_J} \in \mathcal{E}_{N_J}(A)$ *with* $d_{N_J} \notin \mathcal{E}_r(A)$, $r < N_J$, *and* $N_J \to \infty$ *as* $J \to \infty$ *there exists a sequence* $d^*_{N^*_J} \in \mathcal{E}_{N^*_J}[A_0]$ *such that*

$$\overline{\lim_{J\to\infty}} \{ N^*_J / N_J \} \leqq 1 \tag{2.55}$$

and

$$\overline{\lim_{\to\infty}} \{ V_{d^*_{N^*_J}} [t'\hat{\theta}] / F_{d_{N_J}} [t'\hat{\theta}] \} \leqq 1 \qquad \textit{for all } t \in T \tag{2.56}$$

and where ∞ / ∞ *is taken to be* 1.

THEOREM 2.3. *If* A *is a compact convex body in* $E^{(k)}$ *generated by* $b^{(1)}, \cdots, b^{(m)}$, *then* $\mathcal{E}_n[b^{(1)}, \cdots, b^{(m)}]$ *is asymptotically essentially complete* (T) *with respect to* $\mathcal{E}_n[A]$ *for all* $T \subsetneq E^{(k)}$.

PROOF. From theorem 2.2 it is known that for any $d_{N_J} \in \mathcal{E}_{N_J}(A)$ there exists $d^*_{N^*_J}(A_0)$ such that

$$V_{d^*_{N^*_J}} [t'\hat{\theta}] / V_{d_{N_J}} [t'\hat{\theta}] \leqq 1 \tag{2.57}$$

for all $t \in E^{(k)}$ and $N^*_J \leqq N_J + m$. Now since $N_J \to \infty$ we have $\overline{\lim}_{J\to\infty} N^*_J/N_J \leqq 1$.

The following corollaries follow quite readily from theorem 2.3.

COROLLARY 2.3. *With* A *as in corollary* 2.1, $\mathcal{E}_n[b^{(1)}, \cdots, b^{(m)}]$ *is asymptotically essentially complete* (T) *with respect to* $\mathcal{E}_n[A]$ *for all* $T \subsetneq E^{(k)}$.

COROLLARY 2.4. *With* A *as in corollary* 2.2, $\mathcal{E}_n[b^{(1)}, \cdots, b^{(m)}]$ *is asymptotically essentially complete* (T) *with respect to* $\mathcal{E}_n[A]$ *for all* $T \subsetneq E^{(k)}$.

EXAMPLES.

(a) Let us suppose, in this example, that the model is

$$E(Y_J) = \theta_1 x_{J1} + \theta_2 x_{J2} . \tag{2.58}$$

Let the set A be as follows:

$$A = \{ (x_{J1}, x_{J2}) : 0 \leqq a_1 \leqq x_{J1} \leqq b_1;\ 0 \leqq a_2 \leqq x_{J2} \leqq b_2 \} \subsetneq E^{(2)}. \tag{2.59}$$

The set $R(A)$ is the boundary outlined in heavy lines in figure 1, namely, line P_1P_2 and P_2P_3. Theorem 2.1 states that for any n observations in the set A, there exist n ob-

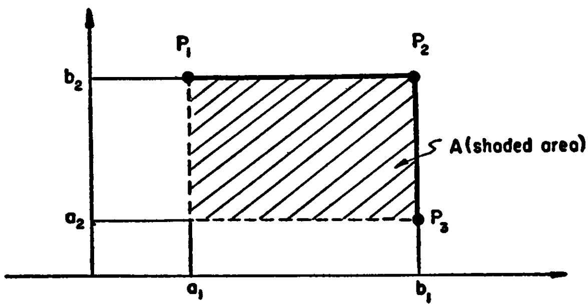

FIGURE 1

servations in $R(A)$ which do as well for estimating any linear combination $t_1\theta_1 + t_2\theta_2$. Corollary 2.2 states that for any n observations in the set A, there exist s observations ($s \leqq n + 3$) at the points P_1, P_2, P_3 that are as efficient for estimating *any* linear combination $t_1\theta_1 + t_2\theta_2$.

(b) Let $E(Y_J) = \theta_1 + \theta_2 x_J$ and $A = \{(1, x_J) | a \leqq x_J \leqq b\} \subsetneq E^{(2)}$. Then A is the heavy line P_1P_2 in figure 2 and $R(A)$ in this case is equal to the set A. Theorem 2.1 is empty for this example. Corollary 2.2 states that for any n observations on the line P_1P_2 there exist s observations ($s \leqq n + 2$) on the points P_1 and P_2 that are as efficient for estimating *any* linear combination $t_1\theta_1 + t_2\theta_2$.

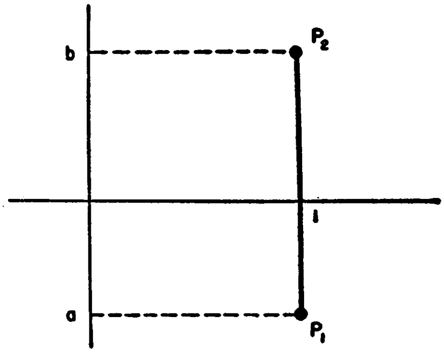

FIGURE 2

(c) Let $Y_{x^{(1)}}, \cdots, Y_{x^{(n)}}$ be n uncorrelated, normally distributed random variables with common variance σ^2.

We suppose

(2.60) $$E(Y_{x^{(J)}}) = \theta' x^{(J)} , \qquad J = 1, \cdots, n .$$

The problem associated with $t = (\theta_{1t}, \theta_{2t})$ will be testing vector θ_{1t} against vector θ_{2t}. The loss functions in this case will be simple. Namely, the loss incurred in problem t when θ_{1t} is chosen is 1 when θ_{2t} is true, and zero otherwise.

The minimax value for problem t when experiment d is used can be calculated as equal to $M(d, t)$ where

(2.61) $$M(d, t) = \Phi\left(\frac{-\Delta_t' F(d) \Delta_t}{\sigma^2}\right),$$

with

(2.62) $$\Delta_t = \theta_{2t} - \theta_{1t} ,$$

(2.63) $$\Phi(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{-u^2/2} du .$$

We first notice that the condition $M(d^*, t) \leqq M(d, t)$ is equivalent to

(2.64) $$\Delta_t' F(d^*) \Delta_t - \Delta_t' F(d) \Delta_t \geqq 0 .$$

It is furthermore noted that (2.64) for all Δ_t is identical with

(2.65) $$V_{d^*}[\Delta_t'\hat{\theta}] \leqq V_d[\Delta_t'\hat{\theta}] \text{ for all } \Delta_t .$$

It becomes clear from this identification that any comparison of minimax values for a problem t can be translated into a comparison of variances of the least squares estimate $\Delta_t'\hat{\theta}$.

Because of the above remark, we can use the proofs of theorem 2.1 and corollary 2.2 to derive the following two theorems.

THEOREM 2.4. *Let $x^{(J)} \in A \subsetneq E^{(k)}$ with A compact. For any $d \in \mathcal{E}_n[A]$ there exists*

$d^* \in \mathcal{E}_n[R(A)]$ such that $M(d^*, t) \leqq M(d, t)$ for all discrimination problems $(\theta_{1t}, \theta_{2t})$ and all n.

THEOREM 2.5. *If the convex closure of $R(A)$ is generated by vector $b^{(1)}, \cdots, b^{(m)}$, then for any $d \in \mathcal{E}_n[A]$ there exists $d^* \in \mathcal{E}_{n+m}[b^{(1)}, \cdots, b^{(m)}]$ such that $M(d^*, t) \leqq M(d, t)$ for all discrimination problems $(\theta_{1t}, \theta_{2t})$ and all n.*

3. Generalization

The model in the previous section will now be generalized. Assume that $Y_{x_1}, \cdots, Y_{x_N}$ are normally distributed (which was not necessary in section 2) with

$$E(Y_{x^{(J)}}) = \psi(\theta, x^{(J)}), \tag{3.1}$$

$$\operatorname{cov}(Y_{x^{(J)}}, Y_{x^{(r)}}) = \begin{cases} \sigma^2 & \text{for } J = r, \\ 0 & \text{for } J \neq r. \end{cases} \tag{3.2}$$

It is assumed that $\partial\psi(\theta, x^{(J)})/\partial\theta_J$ exists for all $\theta \in \Omega$ and all $x^{(J)} \in A$.

When $\psi(\theta, x^{(J)}) = \theta' x^{(J)}$ it is known that the maximum likelihood estimator is the same as the least squares estimator.

The information matrix associated with experiment $d = \mathcal{E}_n(n_1, x^{(1)}; \cdots; n_s, x^{(s)})$ is denoted by $F(d, \theta)$, indicating that it may depend on the values of $\theta \in \Omega$. Now $F(d, \theta)$ is defined to be

$$F(d, \theta) = \sum n_J F_J(d, \theta) \tag{3.3}$$

where

$$F_J(d, \theta) = \left\| \frac{\partial\psi(\theta, x^{(J)})}{\partial\theta_u} \frac{\partial\psi(\theta, x^{(J)})}{\partial\theta_v} \right\|; \qquad u, v = 1, \cdots, m; \tag{3.4}$$

and $\theta' = (\theta_1, \cdots, \theta_m)$ and $x^{(J)\prime} = (x_1^{(J)}, \cdots, x_k^{(J)})$.

It is known that the asymptotic ($n \to \infty$) variance of the maximum likelihood estimate of $\sqrt{n}\, t'\theta = \sqrt{n}(t_1\theta_1 + \cdots + t_m\theta_m)$ is $V_d[\sqrt{n}\, t'\hat{\theta}]$, where

$$V_d[t'\hat{\theta}] = \sigma^2 t' F^{-1}(d, \theta) t \tag{3.5}$$

which now may depend on $\theta \in \Omega$.

In order to adapt the results of section 2, we consider the transformations T_θ from $E^{(k)}$ to $E^{(m)}$

$$T_\theta \begin{pmatrix} x_1^{(J)} \\ \vdots \\ x_k^{(J)} \end{pmatrix} = T_\theta(x^{(J)}) = \begin{pmatrix} \dfrac{\partial\psi(\theta, x^{(J)})}{\partial\theta_1} \\ \vdots \\ \dfrac{\partial\psi(\theta, x^{(J)})}{\partial\theta_m} \end{pmatrix} = \begin{pmatrix} Z_1^{(J)}(\theta) \\ \vdots \\ Z_m^{(J)}(\theta) \end{pmatrix} = Z^{(J)}(\theta). \tag{3.6}$$

It should be noted that when $\psi(\theta, x^{(J)}) = \theta' x^{(J)}$ as in section 2, we have $T_\theta(x) = x$ for all $x \in E^{(k)}$ and all $\theta \in \Omega$.

In terms of the above notation,

$$F_J(d, \theta) = \|Z_u^{(J)}(\theta) Z_v^{(J)}(\theta)\|. \tag{3.7}$$

If the $x^{(J)}$ are restricted to lie in a set $A \subsetneq E^{(k)}$, then for a particular $\theta \in \Omega$, the vector $Z^{(J)}(\theta)$ is restricted to lie in $T_\theta(A)$.

To make $T_\theta(A)$ for all $\theta \in \Omega$ compact it is sufficient that $\partial\psi(\theta, x^{(J)})/\partial\theta_r$ be continuous and bounded for all $\theta \in \Omega$ and $x^{(J)} \in A$. When A is compact, the continuity assumption is sufficient.

Let $A_\theta = \{x^{(J)} \in A \mid T_\theta(x^{(J)}) \in R[T_\theta(A)]\}$ and $A_\Omega = \bigcup_{\theta \in \Omega} A_\theta$.

By use of theorem 2.2 and corollaries 2.1 and 2.2 in section 2 we easily derive theorem 3.1.

THEOREM 3.1. *$\mathcal{E}_n(A_\Omega)$ is asymptotically essentially complete with respect to $\mathcal{E}_n(A)$.*

Let the convex closure of $R[T_\theta(A)]$ be generated by a finite number m_θ of vectors $b^{(1)}(\theta), \cdots, b^{(m)}(\theta)$ for all $\theta \in \Omega$ and let $\tilde{A}_\theta = \{x^{(J)} \in A \mid T_\theta(x^{(J)}) = b^{(s)}(\theta)$ for some $s = 1, \cdots, m\}$ and $\tilde{A}_\Omega = \bigcup_{\theta \in \Omega} \tilde{A}_\theta$. By theorem 2.2 and corollaries 2.1 and 2.2 in section 2 we derive theorem 3.2.

THEOREM 3.2. *$\mathcal{E}_n(\tilde{A}_\Omega)$ is asymptotically essentially complete with respect to $\mathcal{E}_n(A)$.*

EXAMPLES.

(a) Let us suppose $\psi(\theta, x) = \exp(\theta x)$ and $\Omega = \{\theta \mid \theta \geqq 0\} \subsetneq E^{(1)}$ and $0 < x \leqq a < \infty$. The problem $t \in T \in E^{(1)}$ then becomes the problem of estimating $t\theta$, which are all equivalent to estimating θ. We have in this example

$$T_\theta(x) = \frac{\partial \psi(\theta, x)}{\partial \theta} = x e^{\theta x}. \tag{3.8}$$

It is easily computed that

$$A_\Omega = \{a\}, \tilde{A}_\Omega = \{a\}. \tag{3.9}$$

Thus, for large n, observations should be taken at $x = a$. This is reasonable since the regression curves $\exp(\theta x)$ are widest apart at $x = a$.

(b) Let

$$\psi(\theta, x) = \frac{\theta^3}{3} + \frac{(x - \theta)^3}{3}, \qquad 0 \leqq x \leqq 1, \tag{3.10}$$

and $\Omega = \{\theta \mid 0 \leqq \theta \leqq \frac{1}{2}\}$.

Since $R[T_\theta(0 \leqq x \leqq 1)]$ is composed of two points, $\partial\psi(\theta, 0)/\partial\theta$ and $\partial\psi(\theta, 1)/\partial\theta$, we have

$$A_\theta = \{x = 0; x = 1\}. \tag{3.11}$$

Thus

$$A_\Omega = \{0 \leqq x \leqq \tfrac{1}{2};\ x = 1\} \quad \text{and} \quad \tilde{A}_\Omega = A_\Omega. \tag{3.12}$$

There are many questions of interest that have not been fully investigated. Some of these questions are

(1) Is corollary 2.2 the strongest result that can be derived? Under what conditions are the extra m observations not necessary?

(2) What can be said about minimal essentially complete classes?

REFERENCES

[1] HERMAN CHERNOFF, "Locally optimal designs for estimating parameters," *Annals of Math. Stat.*, Vol. 24 (1953), pp. 586–602.

[2] G. ELFVING, "Optimum allocation in linear regression theory," *Annals of Math. Stat.*, Vol. 23 (1952), pp. 255–262.

[3] A. M. MOOD, "On Hotelling's weighing problem," *Annals of Math. Stat.*, Vol. 17 (1946), pp. 432–446.

[4] ABRAHAM WALD, "On efficient design of statistical investigations," *Annals of Math. Stat.*, Vol. 14 (1943), pp. 134–140.

[5] SYLVAIN EHRENFELD, "On the efficiency of experimental designs," *Annals of Math. Stat.*, Vol. 26 (1955), pp. 247–255.