INCOMPLETE MULTIVARIATE DESIGNS, OPTIMAL WITH RESPECT TO FISHER'S INFORMATION MATRIX*

J. N. SRIVASTAVA AND M. K. ZAATAR

(Received Jan. 1, 1972)

Summary

Consider a p-variate normal population $N_p(\boldsymbol{\mu}, \Sigma)$, with $\boldsymbol{\mu}$ known and Σ unknown. Without loss of generality, we take $\boldsymbol{\mu}=\mathbf{0}$. Suppose that we have an incomplete multiresponse sample, i.e., we have samples available from this population and/or its various marginals. Suppose one is interested in estimating Σ, given that all the correlations are known.

Consider the Fisher information matrix H, corresponding to the estimation of the variances σ_{tt}. Consider the marginal involving the responses $i_1, i_2, \cdots, i_k$, and suppose that from this marginal a sample of $n(i_1, i_2, \cdots, i_k)$ is drawn. It is then seen that H is a linear function of the n's. Suppose that the cost of taking an observation on the jth response is ϕ_0, and that a total amount of money ϕ' is available for the collection of samples. The problem considered in this paper is the following. How to choose the n's subject to the cost restriction, such that the determinant of H is maximized. A complete solution is obtained for the case $p=2$. When $p=3$, some partial results are obtained, in particular, it is shown that when all the costs are equal, and the correlations are equal, then the best design is obtained by using a complete sample.

1. Introduction

In most experiments or investigations, usually more than one response or characteristic is measured on an experimental unit or individual. The interest very often lies in estimating certain parameters (usually of the location or scale type) on each response. It is well known that, in general, as the number of units on which a response is observed is increased, the efficiency of the estimate also increases. However, in

* This research was supported by Aerospace Research Laboratories Contract No. F 33615-67-C-1436, Project 7071.

general, one is faced with financial limitations. In other words, there is a certain cost ϕ_j of measuring the jth response on a single unit, and there is a cost ϕ_0 of having a unit available for the experiment. Furthermore, a certain total amount of money (say, 1) is allotted for the experiment. In this situation, not only the number of observations that can be taken on any response is bounded, the number of observations on other responses decreases as the number of observations on a particular response is increased. This indicates that an overall measure of the efficiency of the whole multiresponse experiment is desirable. Also, the number of observations to be taken on the different responses should be determined so as to maximize the overall efficiency of the whole experiment. This is an important aspect of the subject of multiresponse designs.

Certain problems in this field have been studied before. For example, suppose the interest lies in the estimation of the differences between the true effects of a set of v treatments on each of a set of p responses. Thus there will be $v-1$ linearly independent comparisons for each response, giving rise to $p(v-1)$ estimators. Let K denote the $p(v-1)\times p(v-1)$ variance matrix of these estimators. In Srivastava and McDonald [6], the problem of obtaining the optimal multiresponse design (with respect to the minimization of $\operatorname{tr} K$) is considered. This paper is concerned with the case of randomized block designs. A similar paper involving cyclic PBIB designs was considered in Srivastava and McDonald [8]. The first problem under the determinant criterion was considered in Srivastava and McDonald [6].

A multiresponse design problem, posed in a somewhat different fashion has been considered by Hocking and Smith [1].

In this paper, we consider a multiresponse design problem when such parameters are to be estimated. The optimization under the given cost restriction, is done with respect to maximizing the determinant of the fixed information matrix corresponding to these parameters.

2. Preliminaries

Consider the problem of designing an experiment for estimating the unknown dispersion matrix $\Sigma=((\sigma_{ts}))$, $(t, s=1, 2, \cdots, p)$ of a p-variate normal distribution whose mean vector is assumed to be known and equal to zero. (This choice of the known value of the mean causes no loss of generality). The collection of data, for this purpose, could be in the form of a general incomplete multiresponse sample. This sample is divided into subsamples, s_j $(j=1, 2, \cdots, u=2^p-1)$ such that on every experimental unit in S_j, exactly the subset R_j (where $R_j=\{l_{j1}, l_{j2}, \cdots, l_{jp_j}\}$) of responses is measured. Thus u is the number of all non-empty subsets of $\{1, 2, \cdots, p\}$, and n_j is the number of units in S_j. The subsets R_j are assumed to be ordered in any arbitrary but fixed manner. We also assume that the elements of R_j are such that $1 \leqq l_{j1} < l_{j2} < \cdots < l_{jp_j} \leqq p$. The principal submatrix of Σ corresponding to R_j is denoted by Σ_j. For any response r, we define U_r as the union of sets S_j the union being taken over all j such that $r \in R_j$. Thus, U_r is the set of all units on which response r (possibly along with other responses) is measured. We shall also assume a cost restriction of the form mentioned in the introduction, and without loss of generality, label the responses such that $\phi_1 \leqq \phi_2 \leqq \cdots \leqq \phi_p$.

Our objective is to obtain an optimal set of values of the n_j under these cost limitations. By "optimal" we mean the sample should be such as to maximize the "amount of information" obtained from the data within our financial limitations.

In this paper, we choose our measure of information to be the determinant of Fisher's information matrix H. For convenience of reference, we may state here the definition of H. Suppose in any given situation, one is interested in estimating a set of parameters $\theta_1, \cdots, \theta_m$, and suppose L denotes the likelihood of the sample. Then $H=((h_{ij}))$ is $(m \times m)$, and $h_{ij} = \mathrm{E}\{-\partial^2 \log L / \partial\theta_i \partial\theta_j\}$, where E denotes "expectation". Note that the dispersion matrix V of the estimators is asymptotically proportional to the inverse of Fisher's information matrix when these estimators are maximum likelihood estimators, or belong to the class of asymptotically efficient estimators. Also, a plausible property of the determinant criterion is the fact that the $\det H$ and $\det V$ are, respectively, directly and inversely proportional to the volume of the ellipsoid of concentration. See, for example, Roy, Gnanadesiken and Srivastava [5].

A design D, in this study, is a determination of the vector $\boldsymbol{n} = (n_1, n_2, \cdots, n_u)$, specifying, for each subset of responses, the experimental units on which this subset of responses is to be measured. The n_j's must satisfy the cost restriction,

$$1 = \sum_{j=1}^{u} (\zeta_j) n_j \,, \tag{2.1}$$

where $\zeta_j = \phi_0 + \phi_{l_{j1}} + \phi_{l_{j2}} + \cdots + \phi_{l_{jp_j}}$. Under each design D, the criterion of optimization is represented by $[\det H(D)] = Q(D)$, say, where $H(D)$ is the matrix H corresponding to the design D. Thus we are faced with an optimization problem in which Q is the objective function and (2.1) is the constraint. It should be pointed out that the n_j's should also be positive and take on integer values only. However, for practical considerations, the latter requirement shall be overlooked. The cost associated with a design D shall be denoted by $\phi(D)$.

DEFINITION 2.1. A design D is said to be at least as good as a design D', if $Q(D) \geqq Q(D')$ and $\psi(D) \leqq \psi(D')$. Moreover, if any of the two inequalities becomes strict, then D is said to be better and D' is said to be *inadmissible.*

DEFINITION 2.2. *A* (sub) class C of designs is said to be essentially complete in the class C', if $C \subseteq C'$ and for any $D' \in C'$, there exists a $D \in C$, such that D is as good as D'.

3. The case $\Sigma = \theta A$, A known

Here, we shall assume that the positive number θ is the only unknown, and A is a positive definite matrix. An easy computation then gives

$$H = \mathrm{E}\left(-\frac{\partial^2 \log L}{\partial\theta^2}\right) = \frac{\sum_{j=1}^{u} n_j p_j}{2\theta^2}\ . \tag{3.1}$$

Thus we must find n's so as to maximize $\sum_{j=1}^{u} n_j p_j$ subject to (2.1). This is equivalent to maximizing $\Sigma^* n'_j$, subject to $\Sigma^*(\zeta_j/p_j)n'_j = 1$, where Σ^* denotes summation over $j \in J$, where $J = \{j \mid p_j \neq 0\}$. Let j^* be a value of j such that (ζ_j/p_j) attains a minimum (over restriction of j in J) when $j = j^*$. Then clearly, a solution of the problem is: Take $n_j = 0$, if $j \neq j^*$, and take $n_{j^*} = (1/\zeta_{j^*})$. This clearly leads to

THEOREM 3.1. *Let* $\zeta_r^* = (\psi_0 + \psi_1 + \cdots + \psi_r)$; $r = 1, \cdots, p$. *Let* n_r^* *be the size of the sample from the marginal distribution of responses* $1, 2, \cdots, r$. *Let* r^* *be such that* $(r^*/\zeta_{r^*}^*) = \max_{1 \leqq r \leqq p} \{r/\zeta_r^*\}$. *Then the optimal design is obtained by taking a sample of size* $(1/\zeta_{r^*}^*)$ *from the marginal of responses* $1, 2, \cdots, r^*$. *Clearly, the optimal design is an HM design. Also, it is easily checked that this HM design reduces to an SM design if and only if* $\psi_p \leqq \zeta_{p-1}^*/(p-1)$.

4. The case of uncorrelated responses

In the case of uncorrelated responses, the Fisher's information matrix H reduces to a diagonal matrix with $h_{tt} = (2N_t)/\sigma_t^2$, $(t = 1, 2, \cdots, p)$, where $N_t = \sum_{j \in U_t} n_j$. Our problem is to maximize the objective function $Q = \prod_{t=1}^{b} N_t$, subject to the cost restriction (2.1), which is now equivalent to $\psi_0 N_0 + \sum_{t=1}^{b} \psi_t N_t = 1$. The following result can then be proved in a way parallel to Theorem 4.2 in Srivastava and McDonald [7].

THEOREM 4.1. *Let $m_1, \cdots, m_p$ denote respectively the values of $N_1, \cdots, N_p$ which maximizes Q subject to the above cost restriction.*

(a) *Let k be the first integer such that*

$$(4.1) \qquad (k-2)\phi_{k-1} \leqq \zeta^*_{k-2}\,, \qquad (k-1)\phi_k > \zeta^*_{k-1}\,,$$

where $\zeta^*_j = \phi_0 + \phi_1 + \cdots + \phi_j$, and k can take the values $2, 3, \cdots, p$. Then the optimal design is an HM design with

$$(4.2) \qquad m_1 = \cdots = m_{k-1} = (k-1)/(p\zeta^*_{k-1})\,, \quad \text{and} \quad m_j = 1/p\phi_j\,, \quad (j=k, \cdots, p)\,.$$

(b) *If (4.1) does not hold for any value of k $(2 \leqq k \leqq p)$, then we have the SM model with $m_1 = m_2 = \cdots = m_p$.*

5. The case $p=2$, and known correlation

For the case of a bivariate population, let n_1 and n_2 be the sizes of the univariate samples from the marginals of the first and second responses, respectively. The size of the bivariate sample will be denoted by n_3. Also, ρ will denote the correlation coefficient ρ_{12}. The Fisher's information matrix H is given by

$$(5.1) \qquad H = \begin{bmatrix} \dfrac{1}{\sigma_1^2}\left(2n_1 + \dfrac{2-\rho^2}{1-\rho^2}n_3\right) & -\dfrac{1}{\sigma_1\sigma_2}\left(n_3\dfrac{\rho^2}{1-\rho^2}\right) \\ -\dfrac{1}{\sigma_1\sigma_2}\left(n_3\dfrac{\rho^2}{1-\rho^2}\right) & \dfrac{1}{\sigma_2^2}\left(2n_2 + \dfrac{2-\rho^2}{1-\rho^2}n_3\right) \end{bmatrix}.$$

We proceed now to find the optimum design under the determinant criterion. We want to choose n_1, n_2 and n_3 so as to maximize $|H|$ or, equivalently,

$$(5.2) \qquad Q = n_1 n_2 + a(n_1+n_2)n_3 + bn_3^2\,,$$

suject to

$$(5.3) \qquad d = n_1 d_1 + n_2 d_2 + n_3\,,$$

where

$$(5.4) \qquad a = (2-\rho^2)/2(1-\rho^2)\,, \qquad b = 1/(1-\rho^2)\,, \qquad d = 1/(\phi_0+\phi_1+\phi_2)\,,$$
$$d_1 = (\phi_0+\phi_1)/(\phi_0+\phi_1+\phi_2)\,, \qquad d_2 = (\phi_0+\phi_2)/(\phi_0+\phi_1+\phi_2)\,.$$

The inequalities $d_1 \leqq d_2$, $0 \leqq d_1 \leqq 1$, $0 \leqq d_2 \leqq 1$, $d_1 + d_2 \geqq 1$, $a=(b+1)/2$, and $b > a > 1$, can be easily checked.

THEOREM 5.1. *The sub class C^* of HM designs is essentially complete in the class C of GIM designs.*

PROOF. The assumption $\psi_1 \leqq \psi_2$ implies that a design $D(n_1, n_2, n_3)$ in which $n_1 \geqq n_2$ is at least as good as the design $D'(n_2, n_1, n_3)$. Consequently, designs like D' shall be ignored. Thus let $D \in C$. Construct $D^*(m_1, m_2, m_3)$ as follows. Take $m_1 = n_1 - n_2$, $m_2 = 0$, and $m_3 = n_2 + n_3$. The design D^* is of the HM type and an equivalent way of constructing it is to measure the first response on $n_1 + n_3$ experimental units, then measure the second response on a subset of these units of size $n_2 + n_3$. We have

$$Q(D^*) = m_1 m_2 + a(m_1 + m_2)m_3 + bm_3^2 = a(n_1 - n_2)(n_2 + n_3) + b(n_2 + n_3)^2 .$$

It follows that $Q(D^*) - Q(D) = (a-1)n_1 n_2 + (b-a)n_2^2 + 2(b-a)n_2 n_3 \geqq 0$. We also find that $\psi(D) - \psi(D^*) = \psi_0 n_2 > 0$. This completes the proof.

Thus, we will restrict attention to HM designs only, by taking $n_2 = 0$, and $n_1, n_3 \geqq 0$. Our problem is reduced to that of maximizing $Q = an_1 n_3 + bn_3^2$, subject to $d = d_1 n_1 + n_3$, $n_1, n_3 \geqq 0$. It is obvious here that the choice $n_3 = 0$, is never optimal. Now, $n_3 = d - n_1 d_1$, and hence $Q = (d_1 n_1 - d)[(bd_1 - a)n_1 - bd]$. By considering the roots of this quadratic in n_1, one can easily prove

THEOREM 5.2. *A necessary and sufficient condition for the SM model to be optimal is that* $(\psi_2/\psi_0 + \psi_1) \leqq (2+\rho^2)/(2-\rho^2)$. *Otherwise the HM design with* $n_1 = d(2bd_1 - a)/2d_1(bd_1 - a)$ *and* $n_3 = da/2(a - bd_1)$, *is the optimal one.*

6. The case $p=3$

Consider a general incomplete sample from a 3-variate normal population. Let us assume that the variances are unknown, while the (3×3) correlation matrix $((\rho_{rs}))$ $(r, s = 1, 2, 3)$ is known and the mean is equal to the zero vector. For simplicity we shall use the notation: $\rho_{jk} = \rho_i$, where $(i, j, k) \in F$, and F is the set of (the six) permutations of $(1, 2, 3)$. Thus this notation means $\rho_{23} = \rho_{32} = \rho_1$, $\rho_{12} = \rho_{21} = \rho_3$, $\rho_{13} = \rho_{31} = \rho_2$. We find, after some calculations, that the elements h_{ij} of Fisher's information matrix, in this case, are (for $(i, j, k) \in F$) given by

$$(6.1) \quad h_{ii} = \sigma_i^{-2}[2n_i + a_j n_{ij} + a_k n_{ik} + b_i n_{123}] , \qquad h_{ij} = (\sigma_i \sigma_j)^{-1}(C_k n_{123} + d_k n_{ij}) ,$$

where

$$a_i = \frac{2-\rho_i^2}{1-\rho_i^2} , \quad b_i = 1 + \frac{1-\rho_i^2}{\Delta} , \quad c_i = \frac{\rho_1\rho_2\rho_3 - \rho_i^2}{\Delta} , \quad d_i = \frac{-\rho_i^2}{1-\rho_i^2} ,$$

$$\Delta = |((\rho_{ij}))| = 1 - \sum_{i=1}^{3} \rho_i^2 + 2\rho_1\rho_2\rho_3 ,$$

and where n_i $(i = 1, 2, 3)$, n_{ji} $(i \neq j;\ i, j = 1, 2, 3)$ and n_{123} are, respectively, the sample sizes from the corresponding univariate marginals, bivariate marginals and the parent distribution. Define, for $(i, j, k) \in F$,

$$(6.2) \qquad m_i=\sigma_i^2 h_{ii}\,, \quad \mu_k=\sigma_i\sigma_j h_{ij}\,, \quad g_i=\phi_0+\phi_i\,, \quad f=\sum_{r=0}^{3}\phi_r\,, \quad f_i=f-\phi_i\,.$$

Assume that the total money available for taking samples is 1. Our objective is to find a determination of $\boldsymbol{n}'=(n_1, n_2, n_3, n_{12}, n_{13}, n_{23}, n_{123})$ which maximizes $|H|$ or, equivalently, the quantity Q, where

$$(6.3) \qquad Q=\sigma_1^2\sigma_2^2\sigma_3^2|H|=m_1m_2m_3+2\mu_1\mu_2\mu_3-m_1\mu_1^2-m_2\mu_2^2-m_3\mu_3^2\,,$$

subject to the cost restriction

$$(6.4) \qquad 1=g_1n_1+g_2n_2+g_3n_3+f_1n_{23}+f_2n_{13}+f_3n_{12}+fn_{123}\,.$$

A direct attack on this problem, by way of expanding Q as a cubic in the 7 elements of $\boldsymbol{n}$, seems to be unwieldy. The subsequent sections will offer alternative methods of approach.

7. The perturbation method

We start by assuming that the design $D(\boldsymbol{n})$ is optimal, i.e., for any other design $D^+(\boldsymbol{n}^+)$ we have $Q(D^+)\leqq Q(D)$, and $\phi(D)=\phi(D^+)=1$. This means that Q has a global maximum at the point $\boldsymbol{n}$ determined by D. Let us displace the point $\boldsymbol{n}$ to the point $\boldsymbol{n}+\boldsymbol{\delta}$. Suppose this results in a displacement of $(m_1, m_2, m_3, \mu_1, \mu_2, \mu_3)$ by the increments $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \theta_1, \theta_2, \theta_3)$, respectively. The value of Q undergoes the change:

$$(7.1) \qquad \begin{aligned} Q(\boldsymbol{n}+\boldsymbol{\delta})-Q(\boldsymbol{n}) &= \Sigma^*[1/2\varepsilon_i(m_jm_k-\mu_i^2)+\theta_i(\mu_j\mu_k-m_i\mu_i)]+P_0 \\ &= P_1+P_0\,, \qquad \text{say} \end{aligned}$$

where Σ^* shall always denote summation over all triplets $(i, j, k) \in F$, and where P_0 is a polynomial of ε's and θ's of the third degree and

$$(7.2) \qquad \varepsilon_i=2\delta_i+a_k\delta_{ij}+a_j\delta_{ik}+b_i\delta_{123}\,, \qquad \theta_i=d_i\delta_{jk}+c_i\delta_{123}\,.$$

The sign of $Q(\boldsymbol{n}+\boldsymbol{\delta})-Q(\boldsymbol{n})$ is the same as that of P_1, provided that $\boldsymbol{\delta}$ and consequently the ε's and the θ's are such that P_0 is negligible compared to P_1. By the use of (7.2), the polynomial P_1 can be expressed directly in terms of $\boldsymbol{\delta}$ as:

$$(7.3) \qquad P_1=(1/2)\Sigma^*[(\phi_i)\delta_i+\delta_{ij}(a_k\phi_i+a_k\phi_j)2\tau_kd_k)]+\delta_{123}\sum_{i=1}^{3}(b_i\phi_i+2\tau_ic_i)\,,$$

where $\phi_i=m_jm_k-\mu_i^2$, $\tau_i=\mu_k\mu_j-m_i\mu_i$. Because of the cost condition (6.4), the components of $\boldsymbol{\delta}$ satisfy

$$(7.4) \qquad g_1\delta_1+g_2\delta_2+g_3\delta_3+f_3\delta_{12}+f_2\delta_{13}+f_1\delta_{23}+f\delta_{123}=0\,.$$

It is interesting to note that the ϕ's and the τ's are the six different cofactors of the (3×3) matrix $((g_{ij}))$, where $g_{ij}=\sigma_i\sigma_j h_{ij}$.

THEOREM 7.1. *A design $D(\boldsymbol{n})$ is not optimal if there exists a $\boldsymbol{\delta}$ such that $D(n+\boldsymbol{\delta})$ is a design, P_1 is positive and $|P_0|\leqq|P_1|$.*

PROOF. The proof is evident from the preceding development.

This result will be applied to some important special cases.

The SM model

For the SM model we have all the n's equal to zero, except $n_{123}=1/f=m$, say. Consequently, for $(i, j, k)\in F$, we also have $\phi_i=m^2(b_jb_k-c_i^2)$, $\tau_i=(c_jc_k-b_ic_i)m^2$. Thus

(7.5) $\quad P_1=m^2\Sigma^*[\alpha_i\delta_i+(1/2)(a_k\alpha_i+a_k\alpha_j+d_k\beta_k)\delta_{ij}+(1/2)(b_i\alpha_i+c_i\beta_i)\delta_{123}]$,

where

$$\alpha_i=b_jb_k-c_i^2 , \qquad \beta_i=2(c_jc_k-b_ic_i) .$$

Let α^* denote the coefficient of δ_{123} in (7.5), i.e.,

$$\alpha^*=b_1\alpha_1+b_2\alpha_2+b_3\alpha_3+c_1\beta_1+c_2\beta_2+c_3\beta_3 . \tag{7.6}$$

Substituting in (7.5), the value of δ_{123} using the cost condition (7.4), we get

$$P_1=(m^2/2)\Sigma^*[(2d_i-\alpha^*g_i/f)\delta_i+(a_k\alpha_i+a_k\alpha_j+d_k\beta_k-\alpha^*f_k/f)\delta_{ij}] . \tag{7.7}$$

Invoking theorem (6.1) and observing that the δ's in (7.7) must be nonnegative numbers ($\boldsymbol{n}+\boldsymbol{\delta}$ is a design) which can be taken sufficiently small so that $|P_0|<|P_1|$, we obtain

COROLLARY 7.1. *A necessary condition for the SM model to be optimal is that,*

$$\alpha^*>(f)\max_{(i,j,k)\in F}(2\alpha_i/g_i, (a_i\alpha_j+a_i\alpha_k+d_i\beta_i)/f_i) .$$

The GIM model

Consider the GIM design D wit $\boldsymbol{n}>\boldsymbol{0}$, and let η denote the coefficient of δ_{123} in (7.3), i.e.

$$\eta=b_1\phi_1+b_2\phi_2+b_3\phi_3+2\tau_1c_1+2\tau_2c_2+2\tau_3c_3 . \tag{7.8}$$

Substituting for δ_{123} from (7.4) in (7.3), we get

$$P_1=\Sigma^*[(2\phi_i-\eta g_i/f)\delta_i+(a_k\phi_i+a_k\phi_j+2\tau_kd_k-\eta f_k/f)\delta_{ij}](1/2) . \tag{7.9}$$

According to theorem (7.1), a necessary condition for a design D to be optimal, is that P_1 be negative. Moreover, since the δ's can be both negative and positive (the n's>0), their coefficients must vanish. For example, suppose that the coefficient of δ_3, say, is not zero, then we can select all other δ's to be zero, and choose δ_3 to have the same sign as that of its coefficient, thus giving rise to a positive value for P_1, which implies that Q of (6.3) does not have a maximum at the point $\boldsymbol{n}$ determined by D. We have established

COROLLARY 7.2. *A necessary condition for the (GIM) design* $D(n_1>0,\ n_2>0, \cdots, n_{123}>0)$ *to be optimal is that all the following hold simultaneously*

$$\text{(7.10)}\qquad \text{(a)}\quad \eta g_i = 2f\phi_i\ , \qquad \text{(b)}\quad \eta f_i = (a_i\phi_j + a_i\phi_k + 2\tau_1 d_1)f\ .$$

The conditions (7.10), in addition to the cost restriction (6.4) form a set of seven equations in seven unknowns. Thus we could (theoretically) solve for the values of the n's. If a design specifies that some subset of the n's of size x, say, must be zero, then in (7.10) we would have $(6-x)$ equations and x inequalities. Again here, with the cost restriction, we would have to solve $(7-x)$ equations for $(7-x)$ unknowns. For any design D, this procedure determines the optimal values of the positive n's in terms of the costs ϕ_i and the correlation coefficients ρ_i. Any two designs, then, can be compared according to the values of Q arising from adopting each of them.

The HM model

A choice of the vector $\boldsymbol{n}$ such that, say, $n_1>0$, $n_{12}>0$, and $n_{123}>0$, while the rest of n's are equal to zero, gives rise to an HM design. From (6.2) we have $m_1=2n_1+a_3n_{12}+b_1n_{123}$, $m_2=2n_2+a_3n_{12}+b_2n_{123}$, $m_3=b_3n_{123}$, $\mu_1=c_1n_{123}$, $\mu_2+c_2n_{123}$, $\mu_3=d_3n_{12}+c_3n_{123}$. From (7.4) we also get the corresponding values of ϕ_i and τ_i as

$$\phi_1^* = a_3b_3n_{12}n_{123} + (b_2b_3 - c_1^2)n_{123}^2\ ,$$

$$\phi_2^* = b_3(2n_1 + a_3n_{12})n_{123} + (b_1b_3 - c_2^2)n_{123}^2\ ,$$

$$\begin{aligned}\phi_3^* = {}& 2a_3n_1n_{12} + (a_3^2 - d_3^2)n_{12}^2 + [a_3(b_1+b_2) - 2c_3d_3]n_{12}n_{123} + 2b_2n_1n_{123} \\ & + (b_2b_1 - c_3^2)n_{123}^2\ ,\end{aligned}$$

$$\tau_1^* = [(c_2d_3 - a_3c_1)n_{12} - 2c_1n_1]n_{123} + (c_2c_3 - c_1b_1)n_{123}^2\ ,$$

$$\tau_2^* = (c_1d_3 - a_3c_2)n_{12}n_{123} + (c_1c_3 - c_2b_2)n_{123}^2\ ,$$

$$\tau_3^* = (c_1c_2 - c_3b_3)n_{123}^2 - b_3d_3n_{12}n_{123}\ .$$

The quantity η defined in (7.8), after some simplification, takes on the value

$$\eta = 2(a_3b_2b_3 + a_3b_1b_3 - 2b_3c_3d_3 + 2c_1c_2d_3 - a_3c_1^2 - a_3c_2^2)n_{12}n_{123}$$
$$+ 4(b_2b_3 - c_1^2)n_1n_{123} + 2a_3b_3n_1n_{12} + b_3(a_3^2 - d_3^2)n_{12}^2$$
$$+ 3(b_1b_2b_3 + 2c_1c_2c_3 - b_1c_1^2 - b_2c_2^2 - b_3c_3^2)n_{123}^2 .$$

Arguing as before, one arrives at

COROLLARY 7.3. *A necessary condition for the optimality of an HM design with* $n_1, n_{12}, n_{123} > 0$, *is the simultaneous realization of the following*

(a) $\eta g_1 = 2f\phi_1$, (b) $\eta f_3 = (a_3\phi_1 + a_3\phi_2 + 2k_3d_3)f$,
(c) $\eta g_2 \geqq 2f\phi_2$, (d) $\eta g_3 \geqq 2f\phi_3$,
(e) $\eta f_2 \geqq (a_2\phi_1 + a_2\phi_3 + 2k_2d_2)f$,
(f) $\eta f_1 \geqq (a_1\phi_2 + a_1\phi_3 + 2k_1d_1)f$.

Again we remark here that (a) *and* (b) *above with the cost restriction* $1 = g_1n_1 + f_3n_{12} + fn_{123}$ *form a set of three equations from which we find the values of* n_1, n_{12} *and* n_{123}. *These values should satisfy the inequalities* (c), (d), (e) *and* (f).

8. Designs with constant Q

Given a design D, we can find a design D^* that gives the same value for the objective function Q as D does, i.e. $Q(D^*) = Q(D)$, while the cost for D^* is less than that for D. One simple way of achieving this is to define D^* such that the values of the m's and the μ's (see (6.2)) do not change. This is equivalent to saying that the element of $\boldsymbol{\delta}$ satisfy the following conditions, for $(i, j, k) \in F$.

$$(8.1) \qquad 2\delta_i + a_k\delta_{ij} + a_j\delta_{ik} + b_1\delta_{123} = 0 , \qquad d_i\delta_{jk} + c_i\delta_{123} = 0 ,$$
$$2f\delta_{123} + \Sigma^*(g_i\delta_i + f_i\delta_{jk}) < 0 .$$

Obviously, if such $\boldsymbol{\delta}$ exists and the design $D(\boldsymbol{n} + \boldsymbol{\delta})$ has cost 1, then $D(\boldsymbol{n})$ is inadmissible.

From (8.1) we can solve, in terms of δ_{123}, for the rest of the unknowns and obtain

$$(8.2) \qquad \delta_i = (2d_jd_k)^{-1}(a_jc_jd_k + a_kc_kd_j - b_id_jd_k)\delta_{123}$$
$$= \Delta^{-1}[1 + \rho_i^2 - (\rho_i\rho_j/\rho_k) - (\rho_i\rho_k/\rho_j)]\delta_{123}$$
$$\delta_{ij} = -(c_k/d_k)\delta_{123} = [(\rho_1\rho_2\rho_3 - \rho_k^2)/(1 - \rho_k^2)/(\Delta\rho_k^2)]\delta_{123} .$$

Substituting the above values in the inequality (8.1) and multiplying both sides by $2d_1d_2d_3 < 0$, we obtain

$$(8.3) \quad \{\Sigma^*[g_id_i(a_kc_kd_j + a_jc_jd_k - b_id_jd_k)(1/2) - f_ic_id_jd_k] + 2fd_1d_2d_3\}\,\delta_{123} > 0 .$$

Using (6.1), the last inequality is seen to be equivalent to

$$(8.4)\qquad [-(\rho_1\rho_2\rho_3/\Delta)(1/\rho_1^2+1/\rho_2^2+1/\rho_3^2) \\ +(1/2\Delta)(3+\rho_1^2+\rho_2^2+\rho_3^2)-1/2]\delta_{123}\phi_0>0 .$$

By observing that the case $\rho_1\rho_2\rho_3<0$, gives rise to a positive value for the term in brackets in (8.5), and makes $(\delta_1, \delta_2, \delta_3)$ of the same sign as δ_{123}, and $\delta_{12}, \delta_{13}, \delta_{23}$ of a sign opposite to that of δ_{123}, and using a small negative value of δ_{123}, we are led to

THEOREM 8.1. *A design $D(\boldsymbol{n})$ in which n_1, n_2, n_3 and n_{123} are positive, is inadmissible if $\rho_1\rho_2\rho_3<0$. In the following we shall assume that $\rho_1=\rho_2=\rho_3=\rho$, say. For this case, we get*

$$(8.5)\qquad \alpha_i\equiv\alpha=(2-\rho^2)\ ,\qquad \beta_i\equiv\beta=1+(1-\rho^2)/\Delta\ ,$$
$$c_i\equiv c=(\rho^3-\rho^2)/\Delta\ ,\qquad d_i=d=(-\rho^2)/(1-\rho^2)\ ,\qquad (i=1, 2, 3)\ .$$

THEOREM 8.2. *If $\phi_1\leqq\phi_2$ and $\rho_1=\rho_2=\rho_3$, then we must have either $n_1\geqq n_2$, or $n_{13}\geqq n_{23}$, or both.*

PROOF. Suppose the theorem is not true, and let $D(\boldsymbol{n})$ be a design in which $n_1\leqq n_2$ and $n_{13}\leqq n_{23}$. Construct the design $D^*(\boldsymbol{n}^*)$ where $\boldsymbol{n}^*=(n_2, n_1, n_3, n_{12}, n_{13}, n_{23}, n_{123})$. It follows the $\phi(D)-\phi(D^*)=(\phi_2-\phi_1)(n_2-n_1)\geqq 0$, which means that D^* costs less than, or as much as D. Now we shall compare $Q(D)$ with $Q(D^*)$. Let $n_2-n_1=\theta\geqq 0$ and $n_{23}-n_{13}=\theta'\geqq 0$. We find that $m^*=m_1+2\theta$, $m_2^*=m_2-2\theta$, $\mu_1^*=\mu_1$, $\mu_2^*=\mu_2$ and $\mu_3^*=\mu_3$. It can easily be verified that

$$Q(D^*)-Q(D)=2\theta\theta'\{a[2n_3+a(n_{13}+n_{23})+bn_{123})] \\ -d[d(n_{13}+n_{23})+2cn_{123}]\}\ .$$

Thus $Q(D^*)-Q(D)$ is $\geqq 0$ if and only if $2an_3+a^2(n_{13}+n_{23})+abn_{123}\geqq d^2(n_{13}+n_{23})+2dcn_{123}$. From (8.6) we see that $a\geqq 1+|d|$ and $b\geqq 1+|c|$, which implies that the preceding inequality always holds, which in turn implies that $Q(D^*)\geqq Q(D)$. This completes the proof.

9. The case of equal costs and correlation coefficients

In this section we study the case where for all i, we have $\phi_i=\phi$ and $\rho_i=\rho$. We first state a result for later use.

LEMMA 9.1. *Consider a space S of dimension $p(p+1)/2$, whose points are non-negative definite symmetric matrices of size $(p\times p)$. Let $B\in S$ and let $f(B)=|B|^{1/p}$. Then f is a concave function on S.*

PROOF. See, for example, Minc and Marcus ([3], p. 115).

Given $B\in S$, and a permutation α of the elements $\{1, 2, \cdots, p\}$, let

$\alpha(B)$ be a matrix obtained by permuting rows and columns of B (the same permutation for rows as for columns). We have $|\alpha(B)|=|B|$. Let $B^*=1/p! \sum_{\text{all } \alpha} \alpha(B)$.

LEMMA 9.2. *We have* $|B^*|\geqq|B|$.

PROOF. By Lemma 9.1, and the definition of concave function we get

$$|B^*|^{1/p}=\left|\frac{1}{p!}\sum_{\text{all } \alpha}\alpha(B)\right|^{1/p}\geqq\frac{1}{p!}\sum_{\text{all } \alpha}|\alpha(B)|^{1/p}=\frac{1}{p!}p!\,|B|^{1/p},$$

or

$$|B^*|\geqq|B| .$$

Now, for any design $D(n_1, n_2, n_3, n_{12}, n_{13}, n_{23}, n_{123})$, construct $D^*(\nu_1, \nu_1, \nu_1, \nu_2, \nu_2, \nu_2, n_{123})$, where $\nu_1=(n_1+n_2+n_3)/3$, $\nu_2=(n_{12}+n_{13}+n_{23})/3$. Because of equal costs we have $\psi(D^*)=3(\psi_0+\psi)\nu_1+3(\psi_0+2\psi)\nu_2+(\psi_0+3\psi)n_{123}=(\psi_0+\psi)(n_1+n_2+n_3)+(\psi_0+2\psi)(n_{12}+n_{13}+n_{23})+(\psi_0+3\psi)n_{123}=\psi(D)$. And because of equal correlation coefficients we can easily find that $m_1^*=m_2^*=m_3^*=(m_1+m_2+m_3)/3$, and $\mu_1^*=\mu_2^*=\mu_3^*=(\mu_1+\mu_2+\mu_3)/3$. Given any design D, let M be the (3×3) matrix $((m_{ij}))$, with $m_{ii}=m_i$, $m_{12}=\mu_3$, $m_{13}=\mu_2$, and $m_{23}=\mu_1$. It is obvious that $Q(D)=|M|$. It can also be easily checked that $M^*=1/3!\sum_{\alpha=1}^{6}\alpha(M)$. From Lemma 9.1, it follows that $|M^*|\geqq|M|$, or $Q(D^*)\geqq Q(D)$. We have established

THEOREM 9.1. *Consider the case of equal costs and correlations. Given any design* $D(n_1, n_2, n_3, n_{12}, n_{13}, n_{23}, n_{123})$, *the design* $D^*(\nu_1, \nu_1, \nu_1, \nu_2, \nu_2, \nu_2, n_{123})$ *is at least as good as* D, *where* $\nu_1=(n_1+n_2+n_3)/3$, *and* $\nu_2=(n_{12}+n_{13}+n_{23})/3$.

Remark. The extension of this result to general p is immediate.

THEOREM 9.2. *In the case of equal correlations, a design in which* $n_1, n_2, n_3, n_{123}>0$, *is inadmissible.*

PROOF. We shall follow the approach of keeping Q fixed and reducing the cost described in Section 8. From (8.4) we find that the coefficient of ψ_0 is $(1-\rho)/(1+2\rho)$. The fact $\Delta>0$ implies that $\rho>-1/2$, which in turn implies that the coefficient of ψ_0 is always positive. Thus we can take $\delta_{123}<0$. Formulas (8.5) also become

$$\delta_1=\delta_2=\delta_3=(1/1+2\rho)\delta_{123}\ , \qquad \delta_{12}=\delta_{13}=\delta_{23}=-(1+\rho/1+2\rho)\delta_{123}\ .$$

This shows that δ_1, δ_2 and δ_3 are of the same sign as δ_{123}, while δ_{12}, δ_{13} and δ_{23} have a sign opposite to that of δ_{123}. The proof is completed by continuing as in Theorem 8.1.

COROLLARY 9.1. *In the case of equal costs and equal correlations, a design in which both* ν_1 *and* ν_3 *are positive is inadmissible, where* $\nu_1 = n_1 = n_2 = n_3$, and $n_{123} = \nu_3$.

PROOF. Use Theorems 9.1 and 9.2.

LEMMA 9.3. *In the case* $\psi_i = \psi$, $\rho_i = \rho$ $(i = 1, 2, 3)$, *the condition* $\nu_1 = 0$ *implies that the optimal design must have* $\nu_2 = 0$, *and the SM model is optimal.*

PROOF. Let $n_{123} = \nu_3$. Our object is to maximize $Q = m^3 + 2\mu^3 - 3m\mu^2 = (m-\mu)^2(m+2\mu)$, subject to $1 = 3(\psi_0 + 2\psi)\nu_2 + (\psi_0 + 3\psi)\nu_3 = x_1 m + x_2 \mu$, say where $m = 2a\nu_2 + b\nu_3$ and $\mu = d\nu_2 + c\nu_3$. Solving for (ν_2, ν_3) we find

$$\begin{bmatrix} \nu_2 \\ \nu_3 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{2ac - bd} \end{bmatrix} \begin{bmatrix} c & -b \\ -d & 2a \end{bmatrix} \begin{bmatrix} m \\ \mu \end{bmatrix} .$$

Hence

$$x_1 = [(3c-d)\psi_0 + (6c - 3d)\psi]/(2ac - bd) ,$$

$$x_2 = [(2a - 3b)\psi_0 + 6(a-b)\psi]/(2ac - bd) .$$

Now, we have

$$3c - d = -\rho^2(2+\rho)/(1-\rho^2)(1+2\rho) ,$$

$$2c - d = -\rho^2/(1+2\rho)(1-\rho^2) ,$$

$$2a - 3b = 2[(1+\rho)^2 - \rho^3]/(1-\rho^2)(1+2\rho) ,$$

$$a - b = -\rho^2/(1-\rho^2)(1+2\rho) ,$$

$$2ac - bd = -2\rho^2/(1-\rho^2)(1+2\rho) .$$

Hence

$$x_1 = (1/2)(2+\rho)\psi_0 + (3/2) , \qquad x_2 = [((1+\rho)/\rho)^2 - \rho]\psi_0 + 3 .$$

Now $\mu = (1 - x_1 m)/x_2$, and hence $Q(m) = (1/x_2)^3[(x_2 + x_1)m - 1]^2[(x_2 - 2x_1)m + 2]$. This is a cubic polynomial in m, with coefficient of m^3 having the same sign as $(x_2 - 2x_1)(x_2, x_1$ being non-negative). But $x_2 - 2x_1 = [(1-\rho^2)(1+2\rho)/\rho^2]\psi_0 \geqq 0$. The equation $Q(m) = 0$, has a double root m_0 given by $m_0 = [1(x_1 + x_2)] \geqq 0$ $(x_1, x_2 \geqq 0)$. Also, it has a single root $m^+ = [2/(2x_1 - x_2)] \leqq 0$.

In the plane (ν_2, ν_3), the cost constraint restricts the set in which we want to maximize Q, to the line segment between the two points, $[1/(3\psi_0 + 6\psi), 0]$ and $[0, 1/(\psi_0 + 3\psi)]$. This line segment is transformed to a line segment whose projection on the m-axis is the interval (m', m'')

where $m'=a/(3/2\psi_0+3\psi)$, $m''=b/(\psi_0+3\psi)$.

Now $a\leqq b$, which implies that $m'<m''$. At this stage, we would like to show that m' is larger than the value of the double root m_0. Now $m_0=1/[(1/2)(2+\rho)\psi_0+((1+\rho)/\rho)^2\psi_0-\rho\psi_0+a/2\psi]$. Since $a>1$, our objective will be accomplished if we show that $(1/2)(2+\rho)+((1+\rho)/\rho)^2-\rho-3/2\geqq 0$, or $(1+\rho)^2/\rho^2-(1+\rho)/2=(1+\rho)(2+2\rho-\rho^2)/(2\rho^2)=[(1+\rho)(1+2\rho+1-\rho^2)/(2\rho^2)]\geqq 0$. This however, is always true. Thus Q, plotted as a function of m, is tangent to the m axis at m_0, and is a monotone increasing function of m for $m\geqq m_0$. We conclude that Q achieves its restricted maximum at the point $m''=b/(\psi_0+3\psi)$, for which $\mu''=c/(\psi_0+3\psi)$. The inverse image of the point (m'', μ'') in the (ν_2, ν_3) plane is the point $[0, 1/(\psi_0+3\psi)]$. Thus the SM design is the optimal one as the lemma states.

LEMMA 9.4. *In the case of equal costs and correlations, a design in which $\nu_3=0$, cannot be optimal.*

PROOF. Assume that $D(\nu_1, \nu_1, \nu_1, \nu_2, \nu_2, \nu_2, \nu_3)$ is an optimal design in which $\nu_3=0$. We will proceed to determine the exact values of ν_1 and ν_2. We shall also adhere to the same scheme and notation as in the proof of Lemma 9.3. In the present case, we want to maximize $Q=(m-\mu)^2(m+2\mu)$, subject to $1=3(\psi_0+\psi)\nu_1+3(\psi_0+2\psi)\nu_2=x_1m+x_2\mu$, where $m=2\nu_1+2a\nu_2$, $\mu=d\nu_2$. We find that $x_1=(3/2)(\psi_0+\psi)\geqq 0$, and that $x_2=3(1-a)\psi_0/d+3(2-a)\psi/d=3\psi_0/\rho^2+3\psi\geqq 0$. The coefficient of the cubic term in $Q(m)$ has the same sign as x_2-2x_1 which is equal to $(3/\rho^2)\psi_0-3\psi_0\geqq 0$. The equation $Q(m)=0$ has a double root $m_0=[1/(x_1+x_2)]\geqq 0$, and a single root $m^+=[2/(2x_1-x_2)]\leqq 0$. The line segment in the (ν_1, ν_2) plane dictated by the cost restriction is transformed into a line segment whose projection on the m-axis in the (m, μ) plane is the interval (m', m''), where $m'=[2/(3\psi_0+3\psi)]$ and $m''=[2a/(3\psi_0+6\psi)]$. The fact $a-2=[(2-\rho^2)/(1-\rho^2)]-2=\rho^2/(1-\rho^2)\geqq 0$, implies that $m'\leqq m''$. The double root $m_0=[1/(x_1+x_2)]=1/[3\psi_0(1/2+1/\rho^2)+(9/2)\psi]\leqq 1/[\psi_0(3/2)+(3/2)\psi]=m'$. All this leads to the conclusion that $Q(m)$ achieves its restricted maximum at $m''=[2a/(3\psi_0+6\psi)]$, whose inverse image in the (ν_1, ν_2) plane is the point $\nu_1=0$ and $\nu_2=1/(3\psi_0+6\psi)$. Thus the design D which was assumed to be optimal and had $\nu_3=0$, must also have $\nu_1=0$ and $\nu_2=1/(3\psi_0+6\psi)>0$. But in Lemma 9.3 we saw that if $\nu_1=0$ then the optimal design must have $\nu_2=0$. This leads to a contradiction and the proof is complete.

THEOREM 9.3. *In the case of equal costs and equal correlations, the SM design is optimal.*

PROOF. This follows by using successively Theorem 9.1, Corollary 9.1, and Lemmas 9.9 and 9.3.

COLORADO STATE UNIVERSITY

REFERENCES

[1] Hocking, R. R. and Smith, W. B. (1971). Optimal incomplete multinormal samples, To appear in *Technometrics.*

[2] Kleinbaum, D. G. (1970). Estimation and hypothesis testing for generalized multivariate linear models, Unpublished thesis, University of North Carolina, Chapel Hill, N.C.

[3] Marcus, M. and Minc, H. (1964). *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston.

[4] Rao, C. R. (1965). *Linear Statistical Inference and Its Applications*, Wiley and Sons, New York.

[5] Roy, S. N., Gnanadesikan, R. and Srivastava, J. N. (1971). *Analysis and Design of Certain Quantitative Multiresponse Experiments*, Pergamon Press, New York.

[6] Srivastava, J. N. and McDonald, L. L. (1969). On the costwise optimality of hierarchical multiresponse randomized block designs under the trace criterion, *Ann. Inst. Statist. Math.*, **21**, 507-514.

[7] Srivastava, J. N. and McDonald, L. L. (1971). On the cost wise optimality of certain hierarchical and standard multiresponse models under the determinant criterion, *J. Multivariate Analysis*, **1**.

[8] Srivastava, J. N. and McDonald, L. L. (1970). On the hierarchical two-response (cyclic PBIB) designs, costwise optimal under the trace criterion, *Ann. Inst. Statist. Math.*, **22**, 507-518.

[9] Trawinsky, I. M. (1961). Incomplete variable designs, Unpublished thesis, Virginia Polytechnic Institute, Blacksburg, Virginia.