The Annals of Mathematical Statistics
1968, Vol. 39, No. 6, 1978–1994

ON THE ADMISSIBILITY AT ∞, WITHIN THE CLASS OF RANDOMIZED DESIGNS, OF BALANCED DESIGNS[1]

By R. H. Farrell

Cornell University

1. Introduction. The power properties of randomized designs locally about the hypothesis set have been considered by Kiefer [4] and Farrell [1]. Our purpose here is to consider how one should choose a randomized design in order to maximize power far out. This problem is considered when the underlying model is a one way classification and when the underlying model is a two way classification without interaction. Material on the two way classification is to be thought of as being an analysis of a "complex" design problem of interest.

In this section we explain the terminology used. We then state some theorems at the end of this section. The proofs of the theorems require a lemma about convex mixtures of design vectors, proven in Section 2, a lemma about the convexity of the power function of an F-test, proven in Section 3, and estimates on the magnitude of the power of scale invariant tests, obtained in Section 4. Proofs of the theorems are given in subsections of Section 6.

We begin by describing the problem of a two way classification. We assume there are RS populations such that population (r, s) is characterized by a mean value $\varphi_{0r} + \varphi_{1s}$ and let $\varphi_0^T = (\varphi_{01}, \cdots, \varphi_{0R})$, $\varphi_1^T = (\varphi_{11}, \cdots, \varphi_{1S})$, and $\varphi^T = (\varphi_0^T, \varphi_1^T)$. A $H \times 1$ vector X is observed such that $EX = C\varphi$, the matrix $C = (A, B)$ is $H \times (R + S)$ and A is $H \times R$, B is $H \times S$. Each entry in A and B is either 0 or 1, and a given row of A and of B have exactly one non-zero entry. The matrices A and B in effect specify the design and the matrix A^TB has entries n_{rs}, the number of observations on the mean value $\varphi_{0r} + \varphi_{1s}$.

As will appear in the sequel the matrix

$$D = A^TA - (A^TB)(B^TB)^+(B^TA) \tag{1.1}$$

plays an important role in the analysis of variance theory. Corresponding to a square matrix M we let M^+ be its unique generalized inverse in the sense of Penrose [5]. Since the properties of generalized inverses are important to the calculations below we list them here for later reference.

$$\text{(1.2a)} \qquad MM^+M = M; \qquad M^+MM^+ = M^+;$$
$$(M^+M)^T = M^+M; \qquad (MM^+)^T = MM^+.$$

The properties (1.2a) are known to uniquely determine the generalized inverse

Received 28 March 1968; revised 26 July 1968.

[1] This research was supported in part by the Office of Naval Research under Contract Nonr 401(50). Reproduction in whole or in part is permitted for any purpose of the United States Government.

of a square matrix M. If N is any matrix (over the field of real numbers) then

(1.2b) $$N(N^TN)^+(N^TN) = N.$$

Let $D^{\frac{1}{2}}$ be the unique positive semidefinite square root of D. We assume throughout that if I_n is the $n \times n$ identity matrix then

(1.3) $$E(X - C\varphi)(X - C\varphi)^T = \sigma^2 I_H.$$

Then the vector of best linear unbiased estimators

(1.4) $$X_A^* = (D^{\frac{1}{2}})^+(A^T - (A^TB)(B^TB)^+B^T)X$$

has covariance matrix $\sigma^2 DD^+$. From (1.2a) it follows that DD^+ is an orthogonal projection. Therefore if X has a joint normal distribution

(1.5) $$\sigma^{-2}X_A^{*T}X_A^*$$

is a non-central chi-square random variable with centrality (recall that $\varphi^T = (\varphi_0^T, \varphi_1^T)$)

(1.6) $$\varphi_0^T D\varphi_0$$

Note that if e_n is the $n \times 1$ vector such that $e_n^T = (1, 1, \cdots, 1)$,

(1.7) $$De_R = A^T(I_H - B(B^TB)^+B^T)(Ae_R)$$
$$= A^T(I_H - B(B^TB)^+B^T)(Be_S) = 0.$$

Usually the designs used are such that all contrasts of φ_0 are estimable so that D has a single eigenvalue equal zero. Then from (1.6) and (1.7) the statement $\varphi_0^TD\varphi_0 = 0$ is equivalent to $\varphi_{01} = \cdots = \varphi_{0R}$. The standard F-test for the hypothesis $\varphi_0^TD\varphi_0 = 0$ uses $X_A^{*T}X_A^*$ in the numerator.

For reasons that will be apparent in Section 6, it is inconvenient to use X_A^*. Let U_A be a $R \times R$ orthogonal matrix such that

$$U_ADD^+U_A^T = \begin{pmatrix} I_{R-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Then, since D is symmetric, D and D^+ commute, and we may suppose U_A simultaneously diagonalizes D, $D^{\frac{1}{2}}$, and D^+. $EU_AX_A^* = U_AD^{\frac{1}{2}}\varphi_0 = U_AD^{\frac{1}{2}}U_A^TU_A\varphi_0$ so we may write

(1.8) $$(X_A^T, 0) = (U_AX_A^*)^T,$$

and X_A is a $(R - 1) \times 1$ vector with covariance matrix $\sigma^{-2}I_{R-1}$.

In the sequel we will consider only designs for which B^TB is non-singular. Then

(1.9) $$X_B = ((B^TB)^{\frac{1}{2}})^+B^TX$$

is a $S \times 1$ vector with covariance matrix $\sigma^2 I_S$ so that $\sigma^{-2}X_B^TX_B$ is a non-central chi-square random variable. X_A and X_B are stochastically independent and the "sum of squares of error" is

$$Z_C^2 = X^T X - X_A^T X_A - X_B^T X_B . \tag{1.10}$$

In the sequel we let $b > 0$ and define a function φ_b by

$$\text{if} \quad \xi < b \quad \text{then} \quad \varphi_b(\xi) = 0; \tag{1.11}$$
$$\text{if} \quad \xi \geqq b \quad \text{then} \quad \varphi_b(\xi) = 1.$$

In terms of φ_b the standard F-test of $\boldsymbol{\varphi}_0^T D \boldsymbol{\varphi}_0 = 0$ against $\boldsymbol{\varphi}_0^T D \boldsymbol{\varphi}_0 \neq 0$ is to accept the alternative if and only if $\varphi_b(\|X_A\|/Z_C) = 1$.

The total number of observations H satisfies

$$H = \sum_{r=1}^{R} \sum_{r=1}^{S} n_{rs} = e_R^T(A^T B)e_S . \tag{1.12}$$

H is a constant of the problem in that each design considered is required to take H observations. We will suppose T' designs are given. For our analysis it suffices to know the matrices,

$$\text{if} \quad 1 \leqq t \leqq T', \qquad C_t = (A_t, B_t), \quad \text{and} \tag{1.13}$$
$$D_t = A_t^T A_t - (A_t^T B_t)(B_t^T B_t)^+(B_t^T A_t).$$

We will generally speak of $A_t^T B_t$ as the design matrix for the tth design.

Specification of a randomized design consists of specifying three things, the designs, the tests to be used and a probability vector $\mathbf{p}^T = (p_1, \cdots, p_{T'})$ such that the design with matrix C_t is used with probability p_t, $1 \leqq t \leqq T'$. After t is determined, X is observed, $\varphi_t(X)$ is calculated, and the alternative is accepted with conditional probability $\varphi_t(X)$. The power function of C_t, φ_t is

$$\beta(\boldsymbol{\varphi}, \sigma, C_t, \varphi_t) = E_{C_t\boldsymbol{\varphi},\sigma}\varphi_t(X) \tag{1.14}$$
$$= \int (2\pi\sigma^2)^{-H/2}\varphi_t(x) \exp\left(-\|x - C_t\boldsymbol{\varphi}\|^2/2\sigma^2\right) dx.$$

The power function of the *randomized* design is

$$\sum_{t=1}^{T'} p_t \beta(\boldsymbol{\varphi}, \sigma, C_t, \varphi_t). \tag{1.15}$$

The optimality result stated in Theorem 1.1 and proven in Section 6a is about designs for which the matrix D has the special form $D = \alpha_D(I_R - e_R e_R^T/R)$, $\alpha_D > 0$ a constant. It will be recognized that $(I_R - e_R e_R^T/R)$ is a projection matrix of rank $R - 1$ and that e_R is an eigenvector for the eigenvalue 0. A basic problem is to choose design matrices $A^T B$ such that D has the above form and such that $\alpha_D(R - 1) = \text{trace } D$ is maximized. When $H < RS$ so that incomplete block designs are being used the maximization problem is solved by use of a balanced incomplete block design. When $H > RS$ and H is divisible by R any design maximizing trace D gives rise to a matrix $D = \alpha_D(I_R - e_R e_R^T/R)$ (see Theorem 5.5) which is optimal (see Theorem 1.1). The purpose of Section 5 and particularly Theorem 5.5 is to show that for large sample theory Theorem 1.1 is not vacuous.

THEOREM 1.1. *Suppose the design C has matrix $D = \alpha_D(I_R - e_R e_R^T/R)$ and that the design C is a connected design. Suppose the design C maximizes trace D over all possible connected designs taking H observations. Let φ_{1b} be a function of $R + S$ variables such that $\varphi_{1b}(x_A, x_B, z) \equiv \varphi_b(\|x_A\|/z)$. Choose b so that a size α F-test results. Let $C_1, \cdots, C_{T'}, \psi_1, \cdots, \psi_{T'}, p_1, \cdots, p_{T'}$ be a randomized design which is similar size α and such that for some $K_0 > 0$, for all ϱ, σ, if $\|\varrho/\sigma\| > K_0$ then*

$$(1.16) \qquad \beta(\varrho, \sigma, C, \varphi_{1b}) \leqq \sum_{t=1}^{T'} p_t \beta(\varrho, \sigma, C_t, \psi_t).$$

Then

$$(1.17) \qquad D = D_1 = \cdots = D_{T'}.$$

If the functions $\psi_1, \cdots, \psi_{T'}$ are scale change invariant functions of the sufficient statistic then $\psi_1, \cdots, \psi_{T'}$ each agree with the UMP *F-test with probability one.*

The one way classification may be thought of in the preceding terms by introducing a side condition $B = 0$. In this case $D = A^TA$ is usually non-singular and good power properties at ∞ are obtained only if D is non-singular.

For the problem of a one way classification we obtain a somewhat stronger result the statement of which requires further notation. We let Λ be the set of permutations of $1, \cdots, R$ and if $\lambda \varepsilon \Lambda$ we let P_λ be the permutation matrix such that $(P_\lambda^T M P_\lambda)_{ij} = (M)_{\lambda(i)\lambda(j)}$.

THEOREM 1.2. *Let A be the matrix for a design for the one way classification such that if $D = A^TA$ then $|d_{ii} - d_{jj}| \leqq 1$, $1 \leqq i, j \leqq R$, $d_{11}, \cdots, d_{RR}$ the diagonal elements of D. Let φ_b give the* UMP *size α F-test of $\varphi_1 = \cdots = \varphi_R = 0$ and consider the randomized design, use AP_λ, φ_b with probability $(1/R!)$, $\lambda \varepsilon \Lambda$. Let $A_1, \cdots, A_{T'}, \psi_1, \cdots, \psi_{T'}, p_1, \cdots, p_{T'}$ be a randomized design which is similar size α such that for some $K_0 > 0$, for all ϱ, σ, if $\|\varrho\| > \sigma K_0$ then*

$$(1.18) \qquad (R!)^{-1} \sum_{\lambda\varepsilon\Lambda} B(\varrho, \sigma, AP_\lambda, \varphi_b) \leqq \sum_{t=1}^{T'} p_t\beta(\varrho, \sigma, A_t, \psi_t).$$

Then there exist $\lambda_1, \cdots, \lambda_{T'} \varepsilon \Lambda$ such that if $1 \leqq t \leqq T'$, $A^TA = P_{\lambda_t}^T(A_t^TA_t)P_{\lambda_t}$. If each of $\psi_1, \cdots, \psi_{T'}$ are scale change invariant functions then $\psi_1, \cdots, \psi_{T'}$ agree with the UMP *size α F-test with probability one.*

2. Convex mixtures of designs.

LEMMA 2.1. *Let $a^T = (a_1, \cdots, a_R)$ be a vector of non-negative integers such that $a_1 + \cdots + a_R = H$. Let J be the convex hull of the set of vectors $\{P_\lambda^T \mathbf{a}, \lambda \varepsilon \Lambda\}$. Suppose $a_1 \leqq a_2 \leqq \cdots \leqq a_R$ and that $a_1 < a_R$. Then J has dimension $R - 1$ and $(H/R, \cdots, H/R) = (H/R)e_R^T$ is an interior point of J^T in the relative topology on J.*

Suppose $k \geqq 1$ is an integer, that $m \geqq 1$ is an integer, and $H = km + (R - k)\cdot(m + 1)$. Then the vector $\mathbf{a}^*$ defined by $\mathbf{a}^{*T} = (m, \cdots, m, m + 1, \cdots, m + 1)$ is in J. $\mathbf{a}^*$ is an interior point of J in the relative topology of J if and only if $a_1 < m$ and $a_R > m + 1$.

PROOF. The set J lies in the hyperplane $\sum_{i=1}^R x_i = H$. If $\dim J \leqq R - 2$ then there exist constants $(\alpha_1, \cdots, \alpha_R) = \alpha^T$ not all zero and not all pairwise equal such that for all permutations $\lambda \varepsilon \Lambda$, $\alpha^T P_\lambda^T \mathbf{a} =$ constant. By considering interchanges it follows that if $1 \leqq i < j \leqq R$ then $\alpha_i = \alpha_j$. This is a contradiction from which it follows $\dim J = R - 1$.

If $(H/R)e_R$ is a boundary point of J then there exist constants $(\alpha_1, \cdots, \alpha_R) = \alpha^T$ not all zero and a constant such that if $\lambda \varepsilon \Lambda$ then $\alpha^T P_\lambda^T \mathbf{a} \geqq$ constant, for at least one λ, $\alpha^T P_\lambda^T \mathbf{a} >$ constant, and $(H/R)(\alpha^T e_R) =$ constant. This is not possible. Therefore $(H/R)e_R$ is interior to J in the relative topology.

To prove the second part of the lemma, let Λ_1 be the subset of Λ such that if $\lambda \varepsilon \Lambda_1$ then $\lambda(r) = r$, $k + 1 \leqq r \leqq R$. Let Λ_2 be the subset of Λ such that if $\lambda \varepsilon \Lambda_2$ then $\lambda(r) = r$, $1 \leqq r \leqq k$. We let

$$\mathbf{a}_1 = (k!(R-k)!)^{-1} \sum_{\lambda \varepsilon \Lambda_1} \sum_{\tau \varepsilon \Lambda_2} P_\lambda^T P_\tau^T \mathbf{a},$$

and write $\mathbf{a}_1 = (\gamma_1, \cdots, \gamma_1, \gamma_2, \cdots, \gamma_2)$. We now prove $\gamma_1 \leqq m$ and $\gamma_2 \geqq m + 1$. To do this note that $\gamma_1 = (a_1 + \cdots + a_k)/k$ and $\gamma_2 = (a_{k+1} + \cdots + a_R)/(R-k)$. Therefore $a_1 \leqq \gamma_1 \leqq \gamma_2 \leqq a_R$ and $\gamma_1 \leqq a_k \leqq a_{k+1} \leqq \gamma_2$. Thus, if $\gamma_1 > m$ then $a_k > m$ so that since a_k is an integer, $a_k \geqq m + 1$, and it follows that $m + 1 \leqq a_{k+1} \leqq \cdots \leqq a_R$. Hence $H = k\gamma_1 + (R-k)\gamma_2 > km + (R-k) \cdot (m+1) = H$. This contradiction shows $\gamma_1 \leqq m$ must hold. A similar argument shows $\gamma_2 \geqq m + 1$ must hold.

It is clear then that $\mathbf{a}^*$ is a convex mixture of $(H/R)e_R$ and $\mathbf{a}_1$ so that $\mathbf{a}^* \varepsilon J$. Now suppose $a_1 < m$ and $a_R > m + 1$. We show that it follows that $\gamma_1 < m$ and $\gamma_2 > m + 1$. Since $\gamma_1 = (a_1 + \cdots + a_k)/k$ and since $\gamma_1 \leqq m$, if $\gamma_1 = m$ then the assumption $a_1 < m$ implies $a_k \geqq m + 1$. Therefore, since $a_R > m + 1$, $\gamma_2 = (a_{k+1} + \cdots + a_R)/(R-k) > m + 1$, and $H = k\gamma_1 + (R-k)\gamma_2 > km + (R-k)(m+1) = H$. This contradiction shows that it must be the case that $\gamma_1 < m$. A similar argument shows $\gamma_2 > m + 1$. From this it follows that $\mathbf{a}^*$ is a proper convex mixture of $\mathbf{a}_1$ and $(H/R)e_R$. Since $(H/R)e_R$ is an interior point of J in the relative topology it follows that $\mathbf{a}^*$ is an interior point of J in the relative topology.

If $a_1 = m$ or $a_R = m + 1$ then each of the vectors $P_\lambda^T \mathbf{a}$ lies in one of the planes of support $\eta_1(\mathbf{a}') = (a_1' + \cdots + a_k')/k = m$, $\eta_2(\mathbf{a}') = (a_{k+1}' + \cdots + a_R')/(R-k) = m + 1$. □

LEMMA 2.2. *Let M be a $R \times R$ matrix. The matrix $N = (R!)^{-1} \sum_{\lambda \varepsilon \Lambda} P_\lambda^T M P_\lambda$ has the same trace as M, the diagonal elements of N are all equal, and the off diagonal elements of N are all equal.*

PROOF. The ij element of $P_\lambda^T M P_\lambda$ is $(M)_{\lambda(i)\lambda(j)}$. Thus the ii diagonal element is $(R!)^{-1} \sum_{\lambda \varepsilon \Lambda} (M)_{\lambda(i)\lambda(i)} = R^{-1}((M)_{11} + \cdots + (M)_{RR})$. If $i \neq j$ then $\lambda(i) \neq \lambda(j)$. Given $i' \neq j'$ there are exactly $(R-2)!$ permutations λ such that $\lambda(i) = i'$, $\lambda(j) = j'$. Therefore $(N)_{i'j'} = (R(R-1))^{-1} \sum \sum_{i \neq j} (M)_{ij}$. The assertion about the trace of N is obvious.

3. Convexity of the power function of an F-test. The following lemmas are self explanatory and are offered without further comment.

LEMMA 3.1. *Let W be a random variable having a non-central $F_{i,j-i}$ distribution* *and non-centrality parameter* ξ. *There exist constants* c_1 *and* c_2 *depending only on* i *and* j *such that* $EW = c_1 + c_2\xi$, $\xi \geqq 0$.

PROOF. Let $W_{11}, \cdots, W_{1i}$ be independently normally distributed random variables, each with variance equal 1, such that $EW_{11} = \xi^{\frac{1}{2}}$, $EW_{12} = \cdots = EW_{1i} = 0$. Let W_2 be a random variable independent of $W_{11}, \cdots, W_{1i}$ such that W_2 has a chi-square distribution with $j - i$ degrees of freedom. Then

$$EW = EW_2^{-1}(W_{11}^2 + \cdots + W_{1i}^2) \tag{3.1}$$
$$= (i + \xi)EW_2^{-1}.$$

LEMMA 3.2. *Let* $0 \leqq c_3 < c_4$ *and let* $\gamma: [0, \infty] \rightarrow [0, 1]$ *be a Borel measurable function such that if* $x \in [c_3, c_4]$ *then* $\gamma(x) = 1$, *if* $x \notin [c_3, c_4]$ *then* $\gamma(x) = 0$. *Then, if the random variable* W *has a* $F_{i,j-i}$ *distribution, then* $E_{(\cdot)}\gamma(W)$ *has exactly one absolute maximum, say at* ξ_0. *The function* $E_{(\cdot)}\gamma(W)$ *has at most one point of inflexion on the interval* $[\xi_0, \infty]$.

PROOF. If c is a real number the function $\gamma(\cdot) - c$ changes sign at most two times (in the sense of Karlin [2]). Since the family of non-central $F_{i\ j-i}$ distributions are Pólya ∞, see Karlin [2], we may use the sign change properties developed by Karlin [3]. In particular, $E_{(\cdot)}(\gamma(W) - c) = E_{(\cdot)}\gamma(W) - c$ can change sign at most twice. Since $\lim_{\xi\to\infty} E_\xi\gamma(W) = 0$, let the absolute maximum of $E_{(\cdot)}\gamma(W)$ be assumed at ξ_0. If $E_{(\cdot)}\gamma(W)$ has a relative maximum at $\xi_1 \neq \xi_0$ then we consider two cases, $\xi_1 = 0$. In this case there is an $\epsilon > 0$ such that the function $E_{(\cdot)}\gamma(W) - E_0\gamma(W) + \epsilon$ has three zeros, which is impossible. If $\xi_1 \neq 0$ then the slope of $E_{(\cdot)}\gamma(W)$ is zero at ξ_1. Then there must exist $\epsilon > 0$ such that $E_{(\cdot)}\gamma(W) - E_{\xi_1}\gamma(W) + \epsilon$ has two zeros $\xi_{11} < \xi_{12}$ near ξ_1 and a third zero ξ_{13} satisfying $\xi_0 < \xi_{13} < \xi_{11}$. Again this is impossible. Therefore the absolute maximum at ξ_0 is the only relative maximum.

Suppose $E_{(\cdot)}\gamma(W)$ has an inflexion point at $\xi_2 > \xi_0$ and let $E_\xi\gamma(W) = \sum_{i=0}^{\infty} d_i(\xi - \xi_2)^i$. By the preceding paragraph $d_0 \neq 0$ and $d_1 < 0$. We let d_j be the first non-zero coefficient with $j \geqq 3$. In order that the second derivative actually change sign in a neighborhood of ξ_2 requires j to be an odd integer. We consider the function

$$\begin{aligned} E_{(\cdot)}(\gamma(W) - dW - c) &= E_{(\cdot)}\gamma(W) - (dc_2\xi + c + dc_1) \\ &= d_0 - (dc_2\xi_2 + c + dc_1) + (d_1 - dc_2)(\xi - \xi_2) \\ &\quad + d_j(\xi - \xi_2)^j + \text{higher order terms.} \end{aligned} \tag{3.2}$$

We choose as a value of d the value $d = c_2^{-1}(d_1 + \epsilon)$ and let ϵ be a real number of sign the same as d_j such that $d_1 + \epsilon < 0$ and such that the function

$$-\epsilon + d_j(\xi - \xi_2)^{j-1} + \textstyle\sum_{i=j}^{\infty} d_{i+1}(\xi - \xi_2)^i \tag{3.3}$$

has two zeros in a neighborhood of ξ_2. The choice

$$c = d_0 - dc_1 - \xi_2(d_1 + \epsilon) \tag{3.4}$$

makes the constant term of (3.2) equal to zero.

We have shown that at a point ξ_2 such that the second derivative of $E_{(\cdot)}\gamma(W)$ changes sign near ξ_2 we may find a line $dc_2\xi + (c + dc_1)$ which cuts $E_{(\cdot)}\gamma(W)$ at ξ_2 and at two distinct points ξ_{21} and ξ_{22} such that $\xi_{21} < \xi_2 < \xi_{22}$. We observe that the line $dW + c$ can cut $\gamma(W)$ in at most three places so that by Karlin, op. cit., $E_{(\cdot)}\gamma(W) - (dc_2(\cdot) + (c + dc_1))$ can have at most three zeros.

It follows from this analysis that a point ξ_2 of inflexion at which the second derivative of $E_{(\cdot)}\gamma(W)$ changes from plus to minus is not possible. For then if $\xi_{21} < \xi < \xi_2$ we have $E_\xi\gamma(W) < dc_2\xi + c + dc_1$, and if $\xi_2 < \xi < \xi_{22}$ then $E_\xi\gamma(W) > dc_2\xi + c_1 + dc_1$. In a small interval $\xi_{22} < \xi < \xi_{22} + \delta$, $E_\xi\gamma(W) < dc_2\xi + c + dc_1$. Since $\lim_{\xi\to\infty} E_\xi\gamma(W) = 0$, and $\lim_{\xi\to\infty} dc_2\xi + c + dc_1 = -\infty$, it follows there must exist a fourth zero. But such a zero cannot exist.

Let ξ_0 be the location of the absolute maximum of $E_{(\cdot)}\gamma(W)$. Suppose this function has at least two points of inflexion and let $\xi_2 < \xi_3$ be the location of the first two points of inflexion located to the right of ξ_0. By the preceding, the second derivative of $E_{(\cdot)}\gamma(W)$ must be negative on the interval (ξ_0, ξ_2) and therefore there exists $\delta > 0$ such that the second derivative of $E_{(\cdot)}\gamma(W)$ is positive on the interval (ξ_2, ξ_3), negative on the interval $(\xi_3, \xi_3 + \delta)$. This contradiction shows the conclusion of the lemma must hold.

COROLLARY 3.3. *Either $E_{(\cdot)}\gamma(W)$ is a strictly decreasing function or $E_{(\cdot)}\gamma(W)$ has a single inflexion point at $\xi_2 > \xi_0$ such that $E_{(\cdot)}\gamma(W)$ is a strictly convex function on (ξ_2, ∞).*

NOTE. If the derivative of $E_{(\cdot)}\gamma(W)$ vanishes on an interval then $E_{(\cdot)}\gamma(W)$ is a constant function. By completeness, $\gamma(W)$ is a constant with probability one. This contradicts the hypotheses of Lemma 3.2. Therefore $E_{(\cdot)}\gamma(W)$ is strictly convex, $\xi > \xi_2$.

4. Estimates on the power function of scale invariants tests. A function χ of K real variables will be said to be sign and scale change invariant if for all real numbers $x_1, \cdots, x_K$, all choices $\epsilon_1, \cdots, \epsilon_K$ of ± 1, and all real numbers $\delta > 0$, $\chi(\delta\epsilon_1 x_1, \cdots, \delta\epsilon_K x_K) = \chi(x_1, \cdots, x_K)$. χ will be said to be scale change invariant if for all real numbers $\delta > 0$, $x_1, \cdots, x_K$, $\chi(\delta x_1, \cdots, \delta x_K) = \chi(x_1, \cdots, x_K)$.

LEMMA 4.1. *Let W be a $K \times 1$ normally distributed random vector such that $EW = \mu$ and $E(W - \mu)(W - \mu)^T = \epsilon^2 I_K$. Let $Z \geqq 0$ be a random variable independent of W such that Z has a density function $\epsilon^{-(J+1)} c_{1n} z^J \exp(-z^2/2\epsilon^2)$, n an integer, $n \geqq 0$. Let χ be a sign and scale change invariant Borel measurable function of $K + 1$ real variables. Let $a > 0$ be a real number such that if $\|w\|/z > a$ then $\chi(w, z) = 0$. Let $\beta(\mu/\epsilon) = E\chi(W, Z)$. Then, if $\delta' > 0$,*

$$\lim_{\|\mu/\epsilon\|\to\infty} \beta(\mu/\epsilon) \exp(\|\mu/\epsilon\|^2/(2 + 2(a + \delta')^2)) = 0. \tag{4.1}$$

PROOF. The invariance assumption implies there is a Borel measurable function χ^* of K real variables such that for all w and z, $\chi(w, z) = \chi^*(w/z)$. It is therefore sufficient to consider the problem in terms of the joint density function of W/Z which has the form

$$c_{2n} \int_0^\infty z^{K+J} \exp(-\|wz - \mu/\epsilon\|^2/2) \exp(-z^2/2)\, dz. \tag{4.2}$$

Then,

$$(4.3)\quad \beta(\mu/\epsilon) = c_{2n} \int \int_0^\infty \chi^*(w) z^{K+J} \exp(-\|wz - \mu/\epsilon\|^2/2) \exp(-z^2/2)\, dz\, dw.$$

Let $\delta > 0$ and let $c_{3\delta} > 0$ be a constant such that

$$(4.4)\quad \text{if } z > c_{3\delta} \text{ then } z^{K+J} \leqq e^{\delta z^2/2}.$$

Then

$$(4.5)\quad \beta(\mu/\epsilon) \leqq c_{3\delta}^{K+J} c_{2n} \int \int_0^\infty \chi^*(w) \exp(-(\|wz - \mu/\epsilon\|^2 + z^2)/2)\, dz\, dw + c_{2n} \int \int_0^\infty \chi^*(w) \exp(-(\|wz - u/\epsilon\|^2 + (1-\delta)z^2)/2)\, dz\, dw.$$

In each integral of (4.5) substitute $\|wz - \mu/\epsilon\| \geqq |z\|w\| - \|\mu/\epsilon\||$, complete the squares in the two exponents, and extend the range of integration in the resulting integrals to obtain

$$(4.6)\quad \begin{aligned} \beta(\mu/\epsilon) \leqq\ & c_{3\delta}^{K+J} c_{2n} \int \int \chi^*(w) \exp(-(1 + \|w\|^2) \\ & \cdot(z - (\|w\|\,\|\mu/\epsilon\|/(1+\|w\|^2)))^2/2) \\ & \cdot\exp(-\|\mu/\epsilon\|^2/(2 + 2\|w\|^2))\, dz\, dw \\ & + c_{2n} \int \int \chi^*(w) \exp(-((1-\delta) + \|w\|^2) \\ & \cdot(z - (\|w\|\,\|\mu/\epsilon\|/((1-\delta) + \|w\|^2)))^2/2) \\ & \cdot\exp(-\|\mu/\epsilon\|/((2 - 2\delta) + 2\|w\|^2))\, dz\, dw. \end{aligned}$$

The variable z integrates out in (4.6) and we obtain

$$(4.7)\quad \begin{aligned} \beta(u/\epsilon) \leqq\ & c_{3\delta}^{K+J} c_{2n} \int (1 + \|w\|^2)^{-\frac{1}{2}} \chi^*(w) \\ & \cdot\exp(-\|\mu/\epsilon\|/(2 + 2\|w\|^2))\, dw \\ & + c_{2n} \int ((1-\delta) + \|w\|^2)^{-\frac{1}{2}} \chi^*(w) \\ & \cdot\exp(-\|\mu/\epsilon\|/((2 - 2\delta) + 2\|w\|^2))\, dw. \end{aligned}$$

Since χ^* has compact support the integrals in (4.7) are finite. Since $\|w\| > a$ implies $\chi^*(w) = 0$, we find

$$(4.8)\quad \beta(\mu/\epsilon) \leqq c_{4n\delta}(\exp(-\|\mu/\epsilon\|^2/(2 + 2a^2)) + \exp(-\|\mu/\epsilon\|^2/((2 - 2\delta) + 2a^2))).$$

From (4.8) the conclusion of the lemma follows at once.

LEMMA 4.2. *Let W, Z be as for Lemma 4.1. Let χ' be a scale change invariant function of $K + 1$ variables and $\beta(\mu/\epsilon) = E\chi'(W, Z)$ for all μ, $\epsilon > 0$. Suppose for every $\delta > 0$ the set*

$$(4.9)\quad \{(w, z) \mid \chi'(w, z) > 0,\ \|w/z\| \geqq a - \delta\}$$

has positive Lebesgue measure. Then if $\delta' > 0$,

$$(4.10)\quad \lim_{\|\mu/\epsilon\| \to \infty} \beta(\mu/\epsilon) \exp(\|\mu/\epsilon\|^2/(2 + (a - \delta')^2)) = \infty.$$

PROOF. As in the proof of Lemma 4.1, replace χ' by a function χ'^* of K variables. Direct substitution of $\|wz - \mu/\epsilon\| \leqq z\|w\| + \|\mu/\epsilon\|$ into (4.3) shows

$$\begin{aligned}\beta(\mu/\epsilon) \geqq c_{2n} \int \int_0^\infty z^{K+J}\chi'^*(w) \exp\,(-(1 + \|w\|^2) \\ \cdot((z + (\|w\|\,\|\mu/\epsilon\|)/(1 + \|w\|^2)))^2/2) \\ \cdot\exp\,(-\|\mu/\epsilon\|^2/(2 + 2\|w\|^2)).\end{aligned} \tag{4.11}$$

On the set $\|w\| \geqq a - \delta$ we have that

$$\exp\,(-\|\mu/\epsilon\|^2/(2 + 2\|w\|^2)) \geqq \exp\,(-\|\mu/\epsilon\|^2/(2 + 2(a - \delta)^2)). \tag{4.12}$$

Make a change of variable in (4.11) given by,

$$z' = (1 + \|w\|^2)^{\frac{1}{2}}(z + \|\mu/\epsilon\|\cdot\|w\|/(1 + \|w\|^2)), \tag{4.13}$$

and introduce inequality (4.12) into (4.11), to obtain

$$\begin{aligned}\beta(\mu/\epsilon) \geqq c_{2n} \exp\,(-\|\mu/\epsilon\|^2/(2 + 2(a - \delta)^2)) \\ \cdot \int \int_0^\infty \chi_\delta'^*(w)(z(1 + \|w\|^2)^{-\frac{1}{2}} - \|\mu/\epsilon\|\cdot\|w\|(1 + \|w\|^2)^{-1})^{K+J} \\ \cdot\exp\,(-z^2/2)\,dz\,dw.\end{aligned} \tag{4.14}$$

The function $\chi_\delta'^*$ is defined to be

$$\begin{aligned}\chi_\delta'^*(w) &= \chi'^*(w) \quad \text{if} \quad \|w\| \geqq a - \delta; \\ \chi_\delta'^*(w) &= 0 \quad \text{if} \quad \|w\| < a - \delta.\end{aligned} \tag{4.15}$$

The integral in (4.14) is seen to be a non-vanishing polynomial of degree $K + J$ in the variable $\|\mu/\epsilon\|$. The conclusion of Lemma 4.2 now follows from (4.14).

5. Concavity of a design matrix. In Section 1 we introduced the matrix

$$D = A^TA - (A^TB)(B^TB)^+(B^TA).$$

This matrix is symmetric and positive semidefinite so we may consider partial orderings of these matrices by the usual partial ordering of symmetric matrices. A main result in this section is to show that D is a concave function of the matrix A^TB. (Recall that each row of A has a single non-zero entry and that each row of B has a single non-zero entry.) The desired result is stated precisely in Lemma 5.2. From Lemma 5.2 we deduce that a certain class of designs maximize trace D. The result is stated in Theorem 5.5.

LEMMA 5.1. *Let M_1 be an $I \times I$ symmetric matrix, M_2 an $I \times J$ matrix, and M_3 a $J \times J$ symmetric matrix such that $M_2M_3^+M_3 = M_2$. Suppose $M_3 \geqq 0$. Then for all $I \times J$ matrices C,*

$$CM_3C^T + M_2C^T + CM_2^T + M_1 \geqq M_1 - M_2M_3^+M_2^T. \tag{5.1}$$

If $C = -M_2M_3^+$ then equality holds. If $M_3 > 0$ and equality holds then $C = -M_2M_3^+$.

PROOF. Expand the left side of the inequality

$$(C + M_2M_3^+)M_3(C^T + M_3^+M_2^T) \geqq 0. \tag{5.2}$$

Using the identity $M_2M_3^+M_3 = M_2$ we obtain the inequality $CM_3C^T + CM_2^T + M_2C^T + M_2M_3^+M_2^T \geqq 0$. Add M_1 to both sides to obtain (5.1).

It follows from (5.2) that if $C = -M_2M_3^+$ then equality holds. If $M_3 > 0$ then (5.2) can hold if and only if each row vector of $C + M_2M_3^+$ is zero, that is

$$C = -M_2M_3^+. \qquad (5.3)$$ □

The application to be made of Lemma 5.1 uses the identifications $A^TA = M_1$, $A^TB = M_2$, $B^TB = M_3$. By (1.2b) we find

$$M_2M_3^+M_3 = (A^TB)(B^TB)^+(B^TB) = A^TB = M_2. \qquad (5.4)$$

Now suppose if $1 \leqq i \leqq I$ that $C_i = (A_i, B_i)$ is a design, and that $(\alpha_1, \cdots, \alpha_I)$ is a probability vector. Let $N = \sum_{i=1}^I \alpha_i A_i^T B_i$ and assume N is a matrix of integers. Then it is clear that a design $C = (A, B)$ may be chosen such that $A^TB = N$. Further since the non-zero entries of $A_i^TA_i$ and $B_i^TB_i$ are row and column sums of $A_i^TB_i$, $1 \leqq i \leqq I$, it follows that

$$\begin{aligned} A^TA &= \textstyle\sum_{i=1}^I \alpha_i A_i^TA_i; \\ B^TB &= \textstyle\sum_{i=1}^I \alpha_i B_i^TB_i; \\ A^TB &= \textstyle\sum_{i=1}^I \alpha_i A_i^TB_i. \end{aligned} \qquad (5.5)$$

We may therefore calculate that if x is a vector,

$$\begin{aligned} &x^T(A^TA - (A^TB)(B^TB)^+(B^TA))x \\ &= \inf_C x^T \textstyle\sum_{i=1}^I \alpha_i(C(B_i^TB_i)C^T + (A_i^TB_i)C + C^T(B_i^TA_i) + A_i^TA_i)x \\ &\geqq \textstyle\sum_{i=1}^I \alpha_i \inf_C x^T(C(B_i^TB_i)C^T + (A_i^TB_i)C + C^T(B_i^TA_i) + A_i^TA_i)x. \end{aligned} \qquad (5.6)$$

Therefore we have proven

LEMMA 5.2. *Let $C_1, \cdots, C_I$ be designs and $(\alpha_1, \cdots, \alpha_I)$ a probability vector such that $N = \sum_{i=1}^I \alpha_i A_i^TB_i$ is a matrix of integers. If C is any design such that $N = A^TB$ then $D \geqq \sum_{i=1}^I \alpha_i D_i$.* (See (1.1).)

The remainder of this section is devoted to a partially complete characterization of designs C which maximize trace D. The basic formula is

$$\text{if } A^TB = (n_{rs}) \text{ then trace } D = H - \textstyle\sum_{s=1}^S (\sum_{r=1}^R n_{rs}^2/\sum_{r=1}^R n_{rs}). \qquad (5.7)$$

If each integer $n_{rs} = 0$ or $= 1$, then $n_{rs}^2 = n_{rs}$ and trace $D = H - S$.

We now consider the problem of minimizing an expression $(m_1^2 + \cdots + m_R^2)/(m_1 + \cdots + m_R)$. An easy calculation shows that if $j \geqq 2$ then

$$(m + 1)^2 + (m + j - 1)^2 \leqq m^2 + (m + j)^2. \qquad (5.8)$$

Using (5.8) an easy induction shows

LEMMA 5.3. *Subject to $m_1 + \cdots + m_R = m$ the expression $(m_1^2 + \cdots + m_R^2)/m$ is minimized by any vector of integers $(m_1, \cdots, m_R)$ such that if $1 \leqq i, j \leqq R$ then $|m_i - m_j| \leqq 1$.*

LEMMA 5.4. *Let R, a, i be non-negative integers and*

(5.9) $f(a, i) = [(R - i)a^2 + i(a + 1)^2]/[(R - i)a + i(a + 1)]$.

Then $f(a, \cdot)$ is a non-decreasing concave function of i.

PROOF. Treating i as a real variable we find

$$(\partial f/\partial i)(a, i) = (Ra^2 + Ra)/(Ra + i)^2. \qquad \square$$

THEOREM 5.5. *Suppose $H > RS$ and H is divisible by R. Suppose $R((S - i)a + i(a + 1)) = H$. Let C be a design such that $A^T B = (n_{rs})$ satisfies $n_{rs} = a$, $1 \leqq r \leqq R$, $1 \leqq s \leqq S - i$, $n_{rs} = a + 1$, $1 \leqq r \leqq R$, $S - i + 1 \leqq s \leqq S$. Then C maximizes trace D over all possible designs taking H observations.*

PROOF. Let C^* be a design with $A^{*T}B^* = (n^*_{rs})$ such that trace $D^* \geqq$ trace D. In view of Lemma 5.3 we may without loss of generality suppose that if $1 \leqq s \leqq S$, $1 \leqq r_1, r_2 \leqq R$ then $|n_{r_1 s} - n_{r_2 s}| \leqq 1$. Consequently, if $1 \leqq s \leqq S$ we let a_s be the greatest integer such that the entires of the sth column of $A^{*T}B^*$ are equal to a_s or $a_s + 1$. Then if $1 \leqq s \leqq S$,

$$Ra_s \leqq \textstyle\sum_{r=1}^R n^*_{rs} < R(a_s + 1). \tag{5.10}$$

We choose D^* to satisfy

$$\textstyle\sum_{r=1}^R n^*_{r1} \leqq \sum_{r=1}^R n^*_{r2} \leqq \cdots \leqq \sum_{r=1}^R n^*_{rS}.$$

Should it happen that $a_{s_1} = a_{s_2}$ and $\sum_{r=1}^R n^*_{rs_1} = Ra_{s_1} + i_1$, $\sum_{r=1}^R n^*_{rs_2} = Ra_{s_2} + i_2$, $i_1 > 0$ and $i_2 > 0$, then by Lemma 5.4 the trace of D^* may be increased. For suppose $i_1 \geqq i_2$. Then using (5.9)

$$f(a_{s_1}, i_1 + 1) + f(a_{s_1}, i_2 - 1) \leqq f(a_{s_1}, i_1) + f(a_{s_1}, i_2). \tag{5.11}$$

Therefore by induction we may decrease i_2 and increase i_1 by j where j satisfies $i_2 - j = 0$ or $i_1 + j = R$. We may therefore suppose without loss of generality that there is at most one integer s such that $1 \leqq s \leqq S$ and $Ra < \sum_{r=1}^R n^*_{rs} < R(a + 1)$.

We let Λ be the set of permutations of $1, \cdots, R$, Λ_1^* the set of permutations of $1, \cdots, S$ such that $S - i + 1, \cdots, S$ are left fixed, and Λ_2^* the set of permutations of $1, \cdots, S$ such that $1, \cdots, S - i$ are left fixed. Define

$$\begin{aligned} &\text{if} \quad 1 \leqq s \leqq S - i \quad \text{then} \\ &n^{**}_{rs} = (R!(S - i)!)^{-1} \textstyle\sum_{\lambda \varepsilon \Lambda} \sum_{\tau \varepsilon \Lambda_1^*} n^*_{\lambda(r)\tau(s)}; \\ &\text{if} \quad S - i + 1 \leqq s \leqq S \quad \text{then} \\ &n^{**}_{rs} = (R!\, i!)^{-1} \textstyle\sum_{\lambda \varepsilon \Lambda} \sum_{\tau \varepsilon \Lambda_2^*} n^*_{\lambda(r)\tau(s)}. \end{aligned} \tag{5.12}$$

We will now prove

$$\begin{aligned} &\text{if} \quad 1 \leqq s \leqq S - i \quad \text{then} \quad \textstyle\sum_{r=1}^R n^{**}_{rs} \leqq Ra; \\ &\text{if} \quad S - i + 1 \leqq s \leqq S \quad \text{then} \quad \textstyle\sum_{r=1}^R n^{**}_{rs} \geqq R(a + 1). \end{aligned} \tag{5.13}$$

In fact, if $1 \leqq s \leqq S - i$, $\sum_{r=1}^{R} n_{rs}^{**} = (S - i)^{-1} \sum_{s=1}^{S-i} \sum_{r=1}^{R} n_{rs}^{*}$ so that if $\sum_{r=1}^{R} n_{rs}^{**} < R(a + 1)$ for $s \geqq S - i + 1$ then $\sum_{r=1}^{R} n_{r,S-i+1}^{*} < R(a + 1)$ so that $\sum_{r=1}^{R} n_{r,S-i}^{*} \leqq Ra$, and if $1 \leqq s \leqq S - i$ it follows that $\sum_{r=1}^{R} n_{rs}^{**} \leqq Ra$, which is impossible. A similar argument will show from an assumption that if $1 \leqq s \leqq S - i$ and $\sum_{r=1}^{R} n_{rs}^{**} > Ra$ a contradiction results.

If τ is a permutation of $1, \cdots, S$ let Q_τ be the $S \times S$ permutation matrix such that $Q_\tau^T(m_{ij})Q_\tau = (m_{\tau(i)}\ m_{\tau(j)})$. Thus, if Λ^* is the set of permutations of $1, \cdots, S$ we have that (n_{rs}) is in the convex hull of the two matrices

$$(R!S!)^{-1} \sum_{\lambda\varepsilon\Lambda} \sum_{\tau\varepsilon\Lambda^*} P_\lambda^T A^T B\, Q_\tau \qquad \text{and}$$

$$(R!(S - i)!i!)^{-1} \sum_{\lambda\varepsilon\Lambda} \sum_{\tau_1\varepsilon\Lambda_1^*} \sum_{\tau_2\varepsilon\Lambda_2^*} (P_\lambda^T)(A^{*T})B^* Q_{\tau_1} Q_{\tau_2},$$

and hence (n_{rs}) is the convex hull of the matrices

$$P_\lambda^T A^{*T} B^* Q_\tau, \qquad \lambda \varepsilon \Lambda, \qquad \tau \varepsilon \Lambda^*. \tag{5.14}$$

Since each design matrix in (5.14) gives rise to a corresponding matrix

$$P_\lambda^T(A^{*T}A^* - (A^{*T}B^*)(B^{*T}B^*)^+(B^{*T}B^*))P_\lambda \tag{5.15}$$

which has the same trace as D^*, it follows from Lemma 5.2 that

$$\text{trace } D \geqq \text{trace } D^*. \tag{5.16}$$ □

6a. Proof of Theorem 1.1. We assume the design $C = (A, B)$ gives rise to a matrix $D = \alpha_D(I_R - e_R e_R^T/R)$ and that trace $D = \alpha_D(R - 1)$ is maximum among all possible designs taking a total of H observations. We assume a randomized design $C_1, \cdots, C_{T'}, \psi_1, \cdots, \psi_{T'}, p_1, \cdots, p_{T'}$ is given such that (1.15) holds.

From (1.15) we obtain at once that if $\|\varphi\| > \sigma K_0$ then

$$\begin{aligned} 1 - \beta(\varphi, \sigma, C, \varphi_{1b}) &\geqq \textstyle\sum_{t=1}^{T'} p_t(1 - \beta(\varphi, \sigma, C_t, \psi_t)) \\ &\geqq p_t(1 - \beta(\varphi, \sigma, C_t, \psi_t)). \end{aligned} \tag{6a.1}$$

We will use the observation that, since φ_{1b} gives rise to an F-test, $1 - \beta(\varphi, \sigma, C, \varphi_{1b})$ depends only on the parameter $\sigma^{-2}\varphi_0^T D\varphi_0$. As this parameter is scale change invariant, as has been shown in Farrell [1], the test functions $\psi_1, \cdots, \psi_{T'}$ may be replaced by scale change invariant functions $\psi_1^*, \cdots, \psi_{T'}^*$ which are also functions of the sufficient statistics, that is, if $1 \leqq t \leqq T'$, ψ_t^* is a function of $X_{A_t}, X_{B_t}, Z_{C_t}$, and is therefore a function of $R + S$ real variables. The randomized design $C_1, \cdots, C_{T'}, \psi_1^*, \cdots, \psi_{T'}^*, p_1, \cdots, p_{T'}$ is again similar size α and satisfies for all parameter values φ, σ such that $\|\varphi\| > \sigma K_0$,

$$\begin{aligned} 1 - \beta(\varphi, \sigma, C, \varphi_{1b}) &\geqq \textstyle\sum_{t=1}^{T'} p_t(1 - \beta(\varphi, \sigma, C_t, \psi_t^*)) \\ &\geqq p_t(1 - \beta(\varphi, \sigma, C_t, \psi_t^*)). \end{aligned} \tag{6a.2}$$

In order to formulate the problem with care we define for a design $C' = (A', B')$

$$Ex_{A'} = \nu_{A'}, \qquad Ex_{B'} = \nu_{B'}. \tag{6a.3}$$

Then $\nu_{A'}$ is $(R - 1) \times 1$ and $\nu_{B'}$ is $S \times 1$. The $R + S - 1$ parameters represented by $\nu_{A'}$ and $\nu_{B'}$ represent the $R + S - 1$ estimable parametric functions obtainable from $C'^T C' \varphi$. Observe that

$$\nu_{A'}^T \nu_{A'} = \varphi_0^T D' \varphi_0 . \tag{6a.4}$$

The joint density function of $x_{A'}$ and $x_{B'}$ may be written as

$$(2\pi\sigma^2)^{(R+S-1)/2} \exp(-(\|x_{A'} - \nu_{A'}\|^2 + \|x_{B'} - \nu_{B'}\|^2)/2\sigma^2). \tag{6a.5}$$

We suppose that ψ_t^* is then a function of X_{A_t}, X_{B_t} and Z_{C_t} such that ψ_t^* is scale change invariant, $1 \leqq t \leqq T'$.

Let Λ and P_λ be as in Section 1. Since $P_\lambda e_R = e_R$ and since $D_t e_R = 0$, $1 \leqq t \leqq T'$, it follows that $0 = \sum_{\lambda \varepsilon \Lambda} (P_\lambda^T D_\tau P_\lambda) e_R$. Therefore, if $1 \leqq t \leqq T'$,

$$(R!)^{-1} \sum_{\lambda \varepsilon \Lambda} P_\lambda^T D_t P_\lambda = \alpha_{D_t}(I_R - e_R e_R^T/R). \tag{6a.6}$$

Since α_D is maximal,

$$\text{if} \quad 1 \leqq t \leqq T' \quad \text{then} \quad \alpha_{D_t} \leqq \alpha_D . \tag{6a.7}$$

Let $b_t' > 0$ be a number such that

$$\{(x_A, x_B, z) \mid 1 - \psi_t^*(x_A, x_B, z) > 0 \quad \text{and} \quad x_A^T x_A + x_B^T x_B > b_t' z^2\}$$

has positive Lebesgue measure. We apply Lemma 4.2.

$$0 < \inf_{\varphi,\sigma} (1 - \beta(\varphi, \sigma, C_t, \psi_t^*)) \cdot \exp(\sigma^{-2}(\nu_{A_t}^T \nu_{A_t} + \nu_{B_t}^T \nu_{B_t})/(2 + 2(b_t' - \delta)^2)). \tag{6a.8}$$

Thus there exists a constant $c_t > 0$ such that for all parameter values φ, σ satisfying $\|\varphi\| > \sigma K_0$,

$$1 - \beta(\varphi, \sigma, C, \varphi_{1b}) \geqq p_t c_t \exp(-\sigma^{-2}(\varphi_0^T D_t \varphi_0 + \nu_{B_t}^T \nu_{B_t})/(2 + 2(b_t' - \delta)^2)). \tag{6a.9}$$

Therefore for all parameter vectors φ_0/σ such that $\|\varphi_0\| > \sigma K_0$,

$$1 - \beta(\varphi, \sigma, C, \varphi_{1b}) \geqq p_t c_t \exp(-\sigma^2 \varphi_0^T D_t \varphi_0/(2 + 2(b_t' - \delta)^2)), \tag{6a.10}$$

since the left side of (6a.10) is a function of φ through $\varphi_0^{TD}\varphi_0$. We may replace φ_0 by $P_\lambda \varphi_0$ in (6a.10) and sum over $\lambda \varepsilon \Lambda$. Using the convexity of the exponential function and (6a.10),

$$\begin{aligned} &1 - \beta(\varphi, \sigma, C, \varphi_{1b}) \\ &\qquad \geqq p_t c_t \exp(-\sigma^{-2}\alpha_{D_t}(I_R - e_R e_R^T/R)/(2 + 2(b_t' - \delta)^2)) \\ &\qquad = p_t c_t \exp(-\sigma^{-2}(\alpha_{D_t}/\alpha_D)\varphi_0^T D \varphi_0/(2 + 2(b_t' - \delta)^2)). \end{aligned} \tag{6a.11}$$

The inequality (6a.11) is possible if and only if $b_t' \leqq b$, as follows from Lemma 4.1 and the fact that $\beta(\varphi, \sigma, C, \varphi_{1b})$ depends only on $\sigma^{-2}\varphi_0^T D \varphi_0$. If we look at

φ_{1b} as a function of $R + S$ variables x_A, x_B, z then the conclusion just reached says

$$\text{(6a.12)} \qquad \text{if} \quad 1 \leqq t \leqq T', \quad \text{and} \quad (x_A, x_B, z) \varepsilon \mathbb{R}_{R+S},$$

then

$$\psi_t^*(x_A, x_B, z) \leqq \varphi_{1b}(x_A, x_B, z).$$

When $\varphi = 0$, the property of being similar requires

$$\text{(6a.13)} \quad 1 - \alpha = 1 - \beta(0, \sigma, C, \varphi_{1b}) = 1 - \textstyle\sum_{t=1}^{T'} p_t\beta(0, \sigma, C_t, \psi_t^*).$$

If $\varphi = 0$ then $X_{A_1}, \cdots, X_{A_{T'}}$ all have zero means and likewise $X_{B_1}, \cdots, X_{B_{T'}}$ all have zero means. Thus if we let W_A be a normal $(R - 1) \times 1$ vector, W_B be a normal $S \times 1$ vector, $EW_A = 0$, $EW_B = 0$, $EW_AW_A^T = \sigma^2 I_{R-1}$, $EW_BW_B^T = \sigma^2 I_S$, we find

$$\text{(6a.14)} \qquad \begin{aligned} \alpha &= E\varphi_{1b}(W_A, W_B, Z) \\ &= E \textstyle\sum p_t\psi_t^*(W_A, W_B, Z). \end{aligned}$$

In view of (6a.12) the relation (6a.14) can hold only if $\varphi_{1b} = \psi_t^*$ except on a set of Lebesgue measure zero, $1 \leqq t \leqq T'$.

To emphasize the fact that only size α F-tests are being used we write

$$\text{(6a.15)} \qquad \beta(\varphi, \sigma, C, \varphi_{1b}) = \beta^*(\sigma^{-2}\varphi_0^T D\varphi_0);$$

$$\text{if} \quad 1 \leqq t \leqq T', \qquad \beta(\varphi, \sigma, C_t, \psi_t^*) = \beta^*(\sigma^{-2}\varphi_0^T D_t\varphi_0).$$

Then (6a.2) may be written as

$$\text{(6a.16)} \qquad \beta^*(\sigma^{-2}\varphi_0^T D\varphi_0) \leqq \textstyle\sum_{t=1}^{T'} p_t\, \beta^*(\sigma^{-2}\varphi_0^T D_t\varphi_0).$$

By Section 3, let ξ_2 be a real number such that β^* is a concave function on $[\xi_2, \infty)$. From (6a.16) we obtain

$$\text{if} \quad \inf_{1 \leqq t \leqq T', \lambda\varepsilon\Lambda} \sigma^{-2}\varphi_0^T P_\lambda^T D_t P_\lambda\varphi_0 \geqq \xi_2, \quad \text{then}$$

$$\text{(6a.17)} \quad \begin{aligned} \beta^*(\sigma^{-2}\varphi_0^T D\varphi_0) &\leqq \textstyle\sum_{t=1}^{T'} p_t(R!)^{-1} \sum_{\lambda\varepsilon\Lambda} \beta^*(\sigma^{-2}\varphi_0^T P_\lambda^T D_t P_\lambda\varphi_0) \\ &\leqq \textstyle\sum_{t=1}^{T'} p_t\beta^*(\sigma^{-2}\alpha_{D_t}\varphi_0^T(I_R - e_R e_R^T/R)\varphi_0). \end{aligned}$$

Inequality (6a.17) is possible only if $\alpha_{D_t} = \alpha_D$, $1 \leqq t \leqq T'$, and, because β^* is a strictly concave function, for all λ, $P_\lambda^T D_t P_\lambda = D_t$ must hold. This implies $D_t = D$, $1 \leqq t \leqq T'$. □

6b. Proof of Theorem 1.2. For the one way classification the matrix $C = A$ has mutually orthogonal columns, and $D = A^TA$ is a diagonal matrix such that $(D)_{ii}$ is the number of observations on the ith population. We will let $a \geqq 0$ and k, $0 \leqq k < R$, be integers such that $H = (R - k)a + k(a + 1)$ and let D_0 be the special matrix

$$(6b.1)\qquad D_0 = \begin{bmatrix} a & & & & & & & 0 \\ & \cdot & & & & & & \\ & & \cdot & & & & & \\ & & & \cdot & & & & \\ & & & & a & & & \\ & & & & & a+1 & & \\ & & & & & & \cdot & \\ & & & & & & & \cdot \\ 0 & & & & & & & a+1 \end{bmatrix}.$$

In analogy to Section 1, corresponding to a given design A let

$$(6b.2)\qquad X_A = (D^{\frac{1}{2}})^{+}A^{T}X.$$

Then X_A is a random $R \times 1$ vector such that

$$(6b.3)\qquad EX_A = D^{\frac{1}{2}}\varphi; \qquad \operatorname{cov} X_A = \sigma^2 I_R .$$

We let

$$(6b.4)\qquad Z_A = (X^TX - X_A{}^TX_A)^{\frac{1}{2}}.$$

Then the usual F-test of $\varphi = 0$ against $\varphi \neq 0$ is given by, accept the alternative if and only if $\varphi_b(\|X_A\|/Z_A) = 1$. See (1.11).

Let $A_1, \cdots, A_{T'}, \psi_1, \cdots, \psi_{T'}, p_1, \cdots, p_{T'}$ be a similar size α randomized design such that for all parameter values φ, σ satisfying $\|\varphi\| > \sigma K_0$,

$$(6b.5)\qquad (R!)^{-1}\textstyle\sum_{\lambda\varepsilon\Lambda}\beta(\varphi, \sigma, AP_\lambda, \varphi_{1b}) \leqq \sum_{t=1}^{T'} p_t\,\beta(\varphi, \sigma, A_t, \psi_t).$$

If $1 \leqq t \leqq T'$ we let $D_t = A_t{}^TA_t$, a diagonal matrix of integers. In the sequel we shall have occasion to use the lemmas of Section 2 which show $P_\lambda{}^TD_0P_\lambda$ to be in the convex hull of the matrices $\{Q \mid \text{for some } \tau \varepsilon \Lambda,\ Q = P_\tau{}^TD_tP_\tau\}$.

As in the proof of Theorem 1.1 we begin by replacing each of $\psi_1, \cdots, \psi_{T'}$ by scale change invariant functions of the sufficient statistics, $\psi_1^*, \cdots, \psi_{T'}^*$, such that the randomized design $A_1, \cdots, A_{T'}, \psi_1^*, \cdots, \psi_{T'}^*, p_1, \cdots, p_{T'}$ is similar size α and for all parameter values, if $\|\varphi\| > \sigma K_0$ then

$$(6b.6)\qquad (R!)^{-1}\textstyle\sum_{\lambda\varepsilon\Lambda}\beta(\varphi, \sigma, AP_\lambda, \varphi_{1b}) \leqq \sum_{t=1}^{T'} p_t\,\beta(\varphi, \sigma, A_t, \psi_t^*).$$

Then ψ_t^* is a function of $R + 1$ variables, x, z. Let $b' > 0$ be a number such that

$$(6b.7)\qquad \{(x, z) \mid \psi_t^*(x, z) > 0 \quad \text{and} \quad \|x\| > b'z,\ z > 0\}$$

has positive Lebesgue measure. Using Lemma 4.2, if $\delta > 0$ there exists a constant $c_{t\delta}$ such that for all parameter values

$$(6b.8)\qquad 1 - \beta(\varphi, \sigma, A_t, \psi_t^*) \geqq c_{t\delta}\exp(-\sigma^{-2}\varphi^TD_t\varphi/(2 + 2(b' - \delta)^2)).$$

From (6b.6) and (6b.8) we obtain

$$(6b.9)\quad (R!)^{-1} \sum_{\lambda\varepsilon\Lambda} [1 - \beta(\varphi, \sigma, AP_\lambda, \varphi_{1b})] \geqq p_t c_{t\delta}(R!)^{-1} \sum_{\lambda\varepsilon\Lambda} \exp(-\sigma^{-2}\varphi^T P_\lambda^T D_t P_\lambda \varphi/(2 + 2(b' - \delta)^2)) \geqq p_t c_{t\delta}(R!)^{-1} \sum_{\lambda\varepsilon\Lambda} \exp(-\sigma^{-2}\varphi^T P_\lambda^T D_0 P_\lambda \varphi/(2 + 2(b' - \delta)^2)).$$

By Lemma 4.1 we may find a constant c_δ' such that for all parameter values

$$(6b.10)\quad 1 - \beta(\varphi, \sigma, AP_\lambda, \varphi_b) \leqq c_\delta' \exp(-\sigma^{-2}\varphi^T P_\lambda^T D_0 P_\lambda \varphi/(2 + 2(b + \delta)^2)).$$

From (6b.9) and (6b.10) we obtain that for all parameter values, if $\|\varphi\| > \sigma K_0$ then

$$(6b.11)\quad 0 \leqq (R!)^{-1} \sum_{\lambda\varepsilon\Lambda} [c_\delta' \exp(-\sigma^2\varphi^T P_\lambda^T D_0 P_\lambda \varphi/(2 + 2(b + \delta)^2)) - p_t c_{t\delta} \exp(-\sigma^{-2}\varphi^T P_\lambda^T D_0 P_\lambda \varphi/(2 + 2(b' - \delta)^2))].$$

If in (6b.11) we let $\varphi = \xi^{\frac{1}{2}} e_R$ then

$$(6b.12)\quad 0 \leqq c_\delta' \exp(-\sigma^2 \xi H/(2 + 2(b + \delta)^2)) - p_t c_{t\delta} \exp(-\sigma^2 \xi H/(2 + 2(b' - \delta)^2)),$$

to hold for all $\xi > \sigma K_0$. (6b.12) is possible only if $b' \leqq b$.

Therefore if $0 \leqq t \leqq T'$, $\varphi_{1b}(x, z) \geqq \psi_t^*(x, z)$ except possibly on a set of Lebesgue measure zero. Since $c_1, \cdots, c_{T'}, \psi_1^*, \cdots, \psi_{T'}^*, p_1, \cdots, p_{T'}$ is similar size α, and since if $1 \leqq t \leqq T'$, $\beta(0, \sigma, C_t, \psi_t^*)$ does not depend on σ or C_t, it follows as in the proof of Theorem 1.1 that

$$(6b.13)\qquad \text{if}\quad 1 \leqq t \leqq T', \qquad \psi_t^*(\cdot, \cdot) = \varphi_{1b}(\cdot, \cdot)$$

except on a set of Lebesgue measure zero.

Since the functions $\psi_1^*, \cdots, \psi_T^*$ are essentially F-tests,

$$(6b.14)\qquad \text{if}\quad 1 \leqq t \leqq T', \qquad \beta(\varphi, \sigma, C_t, \psi_t^*) = \beta^*(\sigma^{-2}\varphi^T D_t \varphi).$$

From (6b.5) and (6b.14) we obtain

$$(6b.15)\quad (R!)^{-1} \sum_{\lambda\varepsilon\Lambda} \beta^*(\sigma^{-2}\varphi^T P_\lambda^T D_0 P_\lambda \varphi) \leqq (R!)^{-1} \sum_{\lambda\varepsilon\Lambda} \sum_{t=1}^{T'} p_t\, \beta^*(\sigma^{-2}\varphi^T P_\lambda^T D_t P_\lambda \varphi) \leqq (R!)^{-1} \sum_{\lambda\varepsilon\Lambda} \beta^*(\sigma^{-2}\varphi^T P_\lambda^T D_0 P_\lambda \varphi).$$

(6b.15) will hold for all φ/σ such that $\inf_{1 \leqq t \leqq T', \lambda\varepsilon\Lambda} \sigma^{-2}\varphi^T P_\lambda^T D_t P_\lambda \varphi \geqq \xi_2$. Strict inequality will hold unless for all t, $\varphi \neq 0$, and all λ, $\varphi^T D_0 \varphi = \varphi^T P_\lambda^T D_t P_\lambda \varphi$. This implies $D_0 = D_t$, $1 \leqq t \leqq T'$. Theorem 1.2 is therefore proven.

REFERENCES

[1] FARRELL, R. H. (1967). On the admissibility of a randomized symmetrical design for the problem of a one way classification. To appear.

[2] KARLIN, S. (1956). Decision theory for Pólya type distributions. Case of two actions, I. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **1** 125–126. Univ. of California Press.

[3] KARLIN, S. (1957). Pólya type distributions, II. *Ann. Math. Statist.* **28** 281–308.

[4] KIEFER, J. (1958). On the nonrandomized and randomized nonoptimality of symmetrical designs. *Ann. Math. Statist.* **29** 675–699.

[5] PENROSE, R. (1955). A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.* **51** 406–413.