ON THE EFFICIENCY OF EXPERIMENTAL DESIGNS[1]

BY SYLVAIN EHRENFELD

Columbia University

1. Introduction. Many models used in statistical investigations can be formulated in terms of least-square theory. The models that will be discussed can be stated as follows. Let $y_1, \cdots, y_N$ be N independently and normally distributed random variables with common variance σ^2. It is assumed that the expected value of y_α is given by

$$E(y_\alpha) = \beta_1 x_{\alpha 1} + \beta_2 x_{\alpha 2} + \cdots + \beta_p x_{\alpha p}, \qquad \alpha = 1, \cdots, N, \tag{1.1}$$

where quantities $x_{\alpha i}$ for $i = 1, \cdots, p$; $\alpha = 1, \cdots, N$ are known constants and $\beta_1, \cdots, \beta_p$ are unknown constants. The coefficients $\beta_1, \cdots, \beta_p$ are the population regression coefficients of y on $x_1, x_2, \cdots, x_p$ respectively. In matrix notation the above model can be expressed as

$$E(y) = X\beta \tag{1.2}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_p \end{bmatrix}, \qquad X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ x_{N1} & \cdots & x_{Np} \end{bmatrix}, \qquad y = \begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_N \end{bmatrix},$$

The p column vectors in X will be denoted by $x_1, x_2, \cdots, x_p$ where $x_j' = (x_{1j}, x_{2j}, \cdots, x_{Nj})$. In some cases the experimenter has some amount of freedom in the choice of the p vectors x_i. The efficiency and sensitivity of the design may be very much affected by the choice of the design matrix X. The choice of this matrix is equivalent to that of p vectors in N-dimensional Euclidean space.

A simple illustration is furnished by the following example. Suppose y_α are independent random variables with equal variance σ^2, where $E(y_\alpha) = \beta_1(x_\alpha - \bar{x}) + \beta_2$. The x's are assumed to be fixed constants. Suppose, furthermore, that we have N pairs of observations $(x_1, y_1), \cdots, (x_N, y_N)$ and we want to estimate β_1. It is known that the variance of the least square estimate of β_1 is inversely proportional to $\sum_\alpha (x_\alpha - \bar{x})^2$. Hence, if we could choose values $x_1, \cdots, x_N$ in a domain T, we would choose them such that $\sum_\alpha (x_\alpha - \bar{x})$ is as large as possible.

In Section 2 we will prove a theorem about quadratic forms which, with the aid of other considerations, will motivate a criterion for the efficiency of a design matrix X. In Section 3 two theorems will be proved to aid in the applica-

Received July 3, 1953, revised May 7, 1954.

[1] Work partially sponsored by the Office of Naval Research.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

tion of this criterion. In Section 4 applications will be given, and it will furthermore be shown that the Latin square is most efficient in a certain sense.

2. Measure of efficiency of design matrix.

2A. *A useful inequality*. Suppose we have a real $p \times p$ symmetric non-negative matrix S of rank $r \leqq p$. Let t be a column vector with p components, not all zero, such that the equation $S\rho = t$ has solutions for ρ. Then we have the following.

THEOREM 2.1. *If ρ is any solution of $S\rho = t$, then the following inequality holds*:

$$\frac{1}{\lambda_{\max}} \leqq \frac{\rho' S\rho}{t't} \leqq \frac{1}{\lambda_{\min}}, \tag{2.1}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the maximum and the minimum of the nonzero characteristic roots of S.

PROOF. There exists an orthogonal matrix O such that $O'SO = \Omega = (\lambda_i\delta_{ij})$, and therefore $\rho'S\rho = \sum_1^r \lambda_i\rho_i^{*2}$. The $\lambda_1, \cdots, \lambda_r$ are the nonzero characteristic roots of S, and $\rho^* = O'\rho$. Since $S\rho = t$, we also have that $t't = \sum_1^r \lambda_i^2\rho_i^{*2}$ since $t't = \rho^{*\prime}O'S'OO'SO\rho^* = \rho^{*\prime}\Omega^2\rho^*$. Thus if we let $z_i = \sqrt{\lambda_i}\rho_i^*$, we have

$$\frac{\rho'S\rho}{t't} = \frac{\sum_i \lambda_i \rho_i^{*2}}{\sum_i \lambda_i^2 \rho_i^{*2}} = \frac{\sum_i z_i^2}{\sum_i \lambda_i z_i^2}, \tag{2.2}$$

and since

$$\frac{1}{\lambda_{\max}} \leqq \frac{\sum_i z_i^2}{\sum_i \lambda_i z_i^2} \leqq \frac{1}{\lambda_{\min}}, \tag{2.3}$$

we get equation (2.1). When S is of full rank, equation (2.1) becomes

$$\frac{1}{\lambda_{\max}} \leqq \frac{t'S^{-1}t}{t't} \leqq \frac{1}{\lambda_{\min}}. \tag{2.4}$$

For the most part we will restrict ourselves to the case where S has full rank. In that case S has an inverse, and $\rho = S^{-1}t$. The case where S is not of full rank can be treated in a way very similar to the case of full rank. Some of the problems that arise in connection with the model which was considered in Section 1 will be discussed.

2B. *Estimation*. Suppose $S = X'X$ is of full rank, and we want to estimate θ, where $\theta = \sum_1^p t_j\beta_j = t'\beta$ and the t_j are given constants.

From least-square theory it is known that the best unbiased linear estimate of θ, in the sense of minimum variance, is $\hat{\theta} = \sum_j t_j\hat{\beta}_j$ where the $\hat{\beta}_j$ are the least-square estimates of β_j. Also, it is known that the variance of $\hat{\theta}$ is $\sigma^2 t'S^{-1}t$. If t is such that $S\rho = t$ has solutions for ρ, we say that $\theta = t'\beta$ is estimable. Only estimable θ are here considered. When S is not of full rank, the variance of $\hat{\theta}$ is $\sigma^2\rho'S\rho$ where ρ is *any* solution of $S\rho = t$.

2C. *Power*. We will now derive an expression that will be used in Section 4. Suppose S and S^{-1} are partitioned into

$$\left(\begin{array}{c|c} a & c' \\ \hline c & d \end{array}\right) = S, \tag{2.5}$$

$$\left(\begin{array}{c|c} A & B \\ \hline B' & D \end{array}\right) = S^{-1}, \tag{2.6}$$

where a and A are $k \times k$ matrices. If we wanted to test $\beta_1 = \beta_2 = \cdots = \beta_k = 0$, then the usual F test has a power function depending monotonically on a parameter Φ where

$$\sigma^2\Phi = \bar{\beta}'A^{-1}\bar{\beta}, \qquad \bar{\beta}' = (\beta_1, \cdots, \beta_k). \tag{2.7}$$

Since $A^{-1} = a - c'd^{-1}c$, we have

$$\sigma^2\Phi = \bar{\beta}'a\bar{\beta} - (c\bar{\beta})'d^{-1}(c\bar{\beta}). \tag{2.8}$$

2D. *Confidence interval*. If $\sum_j t_j = 0$, then θ is called a contrast. A confidence interval for θ with confidence coefficient $1 - \alpha$ is

$$\hat{\theta} - k_\alpha\hat{\sigma}_{\hat{\theta}} \leqq \theta \leqq \hat{\theta} + k_\alpha\hat{\sigma}_{\hat{\theta}}, \tag{2.9}$$

where k_α is an appropriate constant. Here $\hat{\sigma}_{\hat{\theta}}^2$ is defined by

$$\hat{\sigma}_{\hat{\theta}}^2 = \hat{\sigma}^2t'S^{-1}t, \tag{2.10}$$

where $\hat{\sigma}^2$ is the usual estimate of σ^2. If we let L equal the length of the confidence interval, we have

$$(E(L))^2 = M\sigma^2t'S^{-1}t, \tag{2.11}$$

where M is independent of σ^2.

2E. *Application of inequality*. A simple application of Theorem 2.1 to some of the previous expressions gives

$$\frac{\sigma^2}{\lambda_{\max}}\,t't \leqq \operatorname{Var}(\hat{\theta}) \leqq \frac{\sigma^2}{\lambda_{\min}}\,t't, \tag{2.12}$$

$$\frac{M\sigma^2}{\lambda_{\max}}\,t't \leqq (E(L))^2 \leqq \frac{M\sigma^2}{\lambda_{\min}}\,t't. \tag{2.13}$$

If we wanted to test $\beta_1 = \cdots = \beta_p = 0$, we would get

$$\phi = \frac{\beta'S\beta}{\sigma^2}, \qquad \beta'\beta\,\frac{\lambda_{\max}}{\sigma^2} \geqq \Phi \geqq \frac{\lambda_{\min}}{\sigma^2}\,\beta'\beta. \tag{2.14}$$

It is to be noted that all these bounds can be attained. If (2.12), (2.13), and (2.14) are considered, it can be seen that it would be desirable to make $\lambda_{\min}$ as large as possible. With this in mind, we define the following as a criterion for the efficiency of a design.

DEFINITION. Denote by u the maximum value of $\lambda_{\min}$ for $x_{\alpha i}$ in T. The ratio of $\lambda_{\min}/u$ will be called the *efficiency* of the design. The design will be called *most efficient* if the efficiency of the design is equal to one.

3. Extension. In order to apply this criterion of efficiency, it will be helpful to have the following two theorems.

THEOREM 3.1. *If* $x_i'x_i = c_i$ *where* $c_i \geqq 0$ *are fixed constants, then* $\lambda_{\min}$ *will be a maximum when* $x_i'x_j = 0$ *for* $i \neq j$.

PROOF. Let c be the smallest of $c_1, \cdots, c_p$ and designate

$$S(\lambda) = S - \lambda I, \qquad \bar{S}(\lambda) = \begin{bmatrix} c_1 - \lambda & & 0 \\ & \cdot & \\ & & \cdot \\ & & & \cdot \\ 0 & & c_p - \lambda \end{bmatrix},$$

and let $|S(\lambda)| = |S - \lambda I|$ be the determinant of $S(\lambda)$. We have that $|\bar{S}(\lambda)| = \prod_i(c_i - \lambda)$. Since $S(0)$ and $\bar{S}(0)$ are symmetric and non-negative, we also have $|S(0)| \geqq 0$ and $|\bar{S}(0)| \geqq 0$. Let us now consider the matrix $S(c)$. Suppose first that $S(c)$ is non-negative. We then have $0 \leqq |S(c)| \leqq |\bar{S}(c)| = 0$, and thus c would be a root of $S(0)$ and thus $\lambda_{\min} \leqq c$. Suppose, however, that $S(c)$ were not non-negative. Then there exists a root $\bar{\lambda} \leqq 0$ such that $|S(c) - \bar{\lambda}I| = 0$, or equivalently $|S(0) - cI - \bar{\lambda}I| = |S(0) - (c + \bar{\lambda})I| = 0$. Thus $c + \bar{\lambda}$ is a root of $S(0)$. Since $\bar{\lambda} \leqq 0$, we have $c + \bar{\lambda} \leqq c$ which gives the desired result.

When $x_i'x_j = 0$ for $i \neq j$, x_i is called orthogonal to x_j. When the p vectors are mutually orthogonal, then c, which is the upper bound of $\lambda_{\min}$ can be attained. In some situations one can maintain the condition of orthogonality and increase c at the same time. This theorem is usually useful only if the experimenter has some freedom of choice in all vectors $x_1, x_2, \cdots, x_p$.

In many situations, however, the vectors $x_1, \cdots, x_r$ are fixed, and there is only freedom of choice of $x_{r+1}, \cdots, x_p$. In those cases the following theorem is sometimes useful. Its proof is very similar to the one used for Theorem 3.1.

Suppose $x_1, \cdots, x_r$ are fixed vectors. Then S can be represented as a partitioned matrix as follows:

$$S = S(b, D) = \left(\begin{array}{c|c} A & b \\ \hline b' & D \end{array}\right), \tag{3.1}$$

where $S(b, D)$ indicates that S depends on b and D. Also we have

$$\begin{aligned} A &= (x_i'x_j), & i, j &= 1, \cdots, r, \\ D &= (x_i'x_j), & i, j &= r + 1, \cdots, p. \end{aligned} \tag{3.2}$$

$S(b, D)$, A, and D are symmetric and non-negative matrices. Let the characteristic roots of A be $0 \leqq \lambda_1 \leqq \cdots \leqq \lambda_r$, and those of D be $0 \leqq \theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_{p-r}$. The λ's are fixed, but the θ's depend on the choice of $x_{r+1}, \cdots, x_p$.

THEOREM 3.2. *If ψ is the smallest characteristic root of $S = S(b, D)$, then*

$$\psi \leqq \lambda_1. \tag{3.3}$$

when $b = 0$ and $\lambda_1 \leqq \theta_1$, then $\psi = \lambda_1$.

PROOF. There exist orthogonal matrices T, Q such that

$$T'AT = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{bmatrix}, \quad Q'DQ = \begin{bmatrix} \theta_1 & & 0 \\ & \ddots & \\ 0 & & \theta_{p-r} \end{bmatrix}. \tag{3.4}$$

If we let

$$P = \begin{pmatrix} T & 0 \\ 0 & Q \end{pmatrix}, \tag{3.5}$$

then $P'SP$ has the same characteristic roots as S, since P is an orthogonal matrix. We also get

$$P'SP = \left[\begin{array}{ccc|ccc} \lambda_1 & & 0 & & & \\ & \ddots & & & b^* & \\ 0 & & \lambda_r & & & \\ \hline & & & \theta_1 & & 0 \\ & b^* & & & \ddots & \\ & & & 0 & & \theta_{p-r} \end{array}\right], \qquad b^* = T'bQ. \tag{3.6}$$

Since b^* is zero if and only if b is, we have by use of Theorem 3.1 that $\psi \leqq \min(\lambda_1 \cdots, \lambda_r, \theta_1, \cdots, \theta_{p-r}) \leqq \lambda_1$ which proves the theorem.

In terms of the vectors $x_{r+1}, \cdots, x_p$ Theorem 3.2 means that to maximize $\lambda_{\min}$ we must choose $x_i'x_j = 0$ for $i = 1, \cdots, r$; $j = r + 1, \cdots, p$. In order to attain the upper bound λ_1, we must also make $\lambda_1 \leqq \theta_1$.

To do this we might use Theorem 3.1, and make $x_t'x_s = 0$ for $t, s = r + 1, \cdots, p$ and $x_t'x_t \geqq \lambda_1$ for $t = r + 1, \cdots, p$ if possible.

4. Applications.

4A. *Hotelling's weighing problem.* Consider the weighing of N linear combination of p objects on a chemical balance. This gives rise to the following equations:

$$E(y_\alpha) = x_{\alpha 1}\omega_1 + \cdots + x_{\alpha p}\omega_p, \qquad \alpha = 1, \cdots, N, \tag{4.1}$$

or

$$E(y) = X\omega, \qquad X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{N1} & \cdots & x_{Np} \end{bmatrix} \tag{4.2}$$

where ω_i is the weight of the ith object.

$$x_{\alpha i} = \begin{cases} +1 & \text{if } i\text{th object is placed in left pan at } \alpha\text{th weighing,} \\ -1 & \text{if it is placed in right pan,} \\ 0 & \text{if it is placed in neither.} \end{cases} \tag{4.3}$$

Thus the vectors $x_1, x_2, \cdots, x_p$ are vectors whose components are $+1$, -1, or 0. From Theorem 3.1, it follows that the most efficient design for this problem would be gotten by making $x_i'x_j = 0$ for $i \neq j$ and to make $x_i'x_i$ as large as possible. This can be done by picking $x_{\alpha i} = +1$ or $x_{\alpha i} = -1$, then $x_i'x_i = N$. The vectors $x_1, \cdots, x_p$ cannot be chosen as indicated for all N. The problem of when and how this can be done has been thoroughly discussed in the literature [3].

4B. *Power*. From equation (2.8) it can be seen that $\min_{\beta'\beta \leq 1} \sigma^2\Phi$ would be a maximum if $c = 0$ and the minimum root of a is large. The condition $c = 0$ is equivalent to making $x_i'x_j = 0$ for $i = 1, \cdots, k; j = k + 1, \cdots, p$, and from Theorem 3.1 the $\lambda_{\min}$ of a can be made large by making $x_i'x_i$ for $i = 1, \cdots, k$ as large as possible and $x_s'x_t = 0$ for $s \neq t$; $s, t = 1, \cdots, k$. Any linear hypothesis can be put into canonical form so that it becomes $\beta_1 = \beta_2 = \cdots = \beta_k = 0$.

4C. *Analysis of covariance*. The model usually used for a two-way layout can be stated as follows.

$$E(y_{ij}) = u + Q_i + T_j + \beta(x_{ij} - \bar{x}), \qquad i = 1, \cdots, r; j = 1, \cdots, s, \tag{4.4}$$

where the Q's and T_j's are row and column affects respectively. Associate with each pair (i, j) an integer α such that $\alpha(i, j) = (i - 1)s + j$ for $\alpha = 1, \cdots, rs$. Then (4.4) can be written as

$$E(y_\alpha) = u + \textstyle\sum_i t_{i\alpha} Q_i + \sum_j v_{j\alpha} T_j + \beta(x_\alpha - \bar{x}) \tag{4.5}$$

where

$$\begin{aligned} t_{i\alpha} &= +1 \text{ if } \alpha\text{th plot is in } i\text{th row and 0 otherwise,} \\ v_{j\alpha} &= +1 \text{ if } \alpha\text{th plot is in } j\text{th column and 0 otherwise.} \end{aligned} \tag{4.6}$$

In matrix notation the above becomes

$$E(y) = X\,\theta = (1 \;\; t_{1\alpha} \;\; t_{2\alpha} \cdots t_{r\alpha} \;\; v_{1\alpha} \cdots v_{s\alpha} \;\; (x_\alpha - \bar{x})) \begin{bmatrix} u \\ Q_1 \\ \vdots \\ Q_r \\ T_1 \\ \vdots \\ T_s \\ \beta \end{bmatrix}. \tag{4.7}$$

It can be seen that $x_1, \cdots, x_{r+1}, x_{r+2}, \cdots, x_{r+s+1}$ are fixed vectors in the model. The only choice available to the experimenter is in the last vector x_{r+s+2}.

An application of Theorem 3.2 gives us the following advice for getting a good design. Make $x'_{r+s+2}x_{r+s+2}$ as large as possible, where

$$x'_{r+s+2}x_{r+s+2} = \sum_\alpha (x_\alpha - \bar{x})^2 = \sum_{i,j} (x_{ij} - \bar{x})^2. \tag{4.8}$$

Furthermore, it is desirable to make $x'_{r+s+2}x_j = 0$ for $j = 1, \cdots, r + s + 1$ simultaneously, that is;

$$\sum_\alpha t_{i\alpha}(x_\alpha - \bar{x}) = \sum_j (x_{ij} - \bar{x}) = 0, \tag{4.9}$$

$$\sum_\alpha v_{j\alpha}(x_\alpha - \bar{x}) = \sum_i (x_{ij} - \bar{x}) = 0. \tag{4.10}$$

Equations (4.9) and (4.10) thus give conditions for an efficient design for this model.

4D. *Latin square.* Suppose we wish to find out by experimentation whether there is any significant difference among yields of m different varieties $v_1, \cdots, v_m$. An experimental area is divided into m^2 plots lying in m rows and m columns and each plot is assigned to one of the varieties $v_1, \cdots, v_m$. Denote by y_{ijk} the yield of variety v_k on plot in ith row and jth column. It is assumed that the y_{ijk} are independently and normally distributed with common variance σ^2.

$$E(y_{ijk}) = k_i + d_j + \rho_k \tag{4.11}$$

where k_i, d_j, ρ_k are the row, column, and variety effects respectively. The quantities σ^2, k_i, d_j, ρ_k are unknown parameters. The hypothesis to be tested is that variety has no effect on yield, i.e.,

$$\rho_1 = \rho_2 = \cdots = \rho_k. \tag{4.12}$$

Wald proved that the Latin square was most efficient in a sense which he defined in [1]. We will prove a similar result in this section.

THEOREM 4.1. *The Latin square is most efficient in the sense of our definition.*

PROOF. As was done in (4B) we let $\alpha(i, j) = (i - 1)m + j$ for $i, j = 1, \cdots, m$. Let $t_{i\alpha}$, $u_{j\alpha}$, $z_{k\alpha}$ for $i, j, k = 1, \cdots, m$; $\alpha = 1, \cdots, m^2$ be defined as follows.

$$\begin{aligned} t_{i\alpha} &= +1 \text{ if } \alpha\text{th plot in } i\text{th row and 0 otherwise,} \\ u_{j\alpha} &= +1 \text{ if } \alpha\text{th plot in } j\text{th column and 0 otherwise,} \\ z_{k\alpha} &= +1 \text{ if } k\text{th variety is assigned to } \alpha\text{th plot.} \end{aligned} \tag{4.13}$$

Then (4.11) can be rewritten as follows.

$$E(y_\alpha) = \sum_{i=1}^m k_i t_{i\alpha} + \sum_{j=1}^m d_j u_{j\alpha} + \sum_{k=1}^m \rho_k z_{k\alpha}. \tag{4.14}$$

We want to express the hypothesis in (4.12) in canonical form. In order to do this denote the arithmetic means

$$\bar{t}_1 = \frac{1}{m^2}\sum_\alpha t_{i\alpha}; \qquad \bar{u}_i = \frac{1}{m^2}\sum_\alpha u_{i\alpha}; \qquad \bar{z}_i = \frac{1}{m^2}\sum_\alpha z_{i\alpha} \tag{4.15}$$

and then let $t'_{i\alpha} = t_{i\alpha} - \bar{t}_i$, $u'_{i\alpha} = u_{i\alpha} - \bar{u}_i$, $z'_{i\alpha} = z_{i\alpha} - \bar{z}_i$ and $k'_i = k_i - k_m$, $d'_i = d_i - d_m$, $\rho'_i = \rho_i - \rho_m$ for $i = 1, \cdots, m-1$, and let $\omega_\alpha = 1$ for $\alpha = 1, \cdots, m^2$. Then (4.14) can be written as

$$E(y_\alpha) = \xi\,\omega_\alpha + \sum_{i=1}^{m-1} k'_i t'_{i\alpha} + \sum_{i=1}^{m-1} d'_i u'_{i\alpha} + \sum_{i=1}^{m-1} \rho'_i z'_{i\alpha}, \tag{4.16}$$

where $\xi = \sum_1^{m-1} k'_i \bar{t}_i + \sum_1^{m-1} d'_i \bar{u}_i + \sum_1^{m-1} \rho'_i \bar{z}_i + k_m + d_m + \rho_m$, and the hypothesis in canonical form is

$$\rho'_1 = \rho'_2 = \cdots = \rho'_{m-1} = 0. \tag{4.17}$$

In matrix notation the design matrix X becomes

$$X = (1 \quad t'_{1\alpha} \quad t'_{z\alpha} \cdots t'_{m-1,\alpha} \quad u'_{1\alpha} \cdots u'_{m-1,\alpha} \quad z'_{1\alpha} \cdots z^{1}_{m-1,\alpha}). \tag{4.18}$$

The experimenter has freedom only in the choice of the $z'_{i\alpha}$, $i = 1, \cdots, m-1$; $\alpha = 1, \cdots, m^2$. The $z'_{i\alpha}$ depend on the way the varieties $v_1, \cdots, v_m$ are assigned to the m^2 plots. In the Latin square arrangement each variety appears exactly once in each row and exactly once in each column. The S matrix for the above model is

$$S = \left[\begin{array}{c|ccc|ccc|c} m^2 & & 0 & & & 0 & & \\ \hline & m-1 & & -1 & & & & \\ 0 & & \ddots & & & 0 & & \\ & -1 & & m-1 & & & & b \\ \hline & & & & m-1 & & -1 & \\ 0 & & 0 & & & \ddots & & \\ & & & & -1 & & m-1 & \\ \hline & & & b' & & & & D \end{array}\right] \tag{4.19}$$

and we let $D = (z''_i z'_j)$ and

$$A = \left[\begin{array}{c|ccc|ccc} m^2 & & 0 & & & 0 & \\ \hline & m-1 & & -1 & & & \\ 0 & & \ddots & & & 0 & \\ & -1 & & m-1 & & & \\ \hline & & & & m-1 & & -1 \\ 0 & & 0 & & & \ddots & \\ & & & & -1 & & m-1 \end{array}\right]. \tag{4.20}$$

It is known that for the Latin square

$$\textstyle\sum_\alpha z'_{k\alpha} u'_{j\alpha} = \sum_\alpha z'_{k\alpha} t'_{i\alpha} = \sum_\alpha z'_{k\alpha}\omega_\alpha = 0, \qquad i, j, k = 1, \cdots, m-1, \tag{4.21}$$

which is exactly the condition of orthogonality needed in Theorem 3.2, namely $b = 0$. For a Latin square the D matrix becomes

$$D = \begin{bmatrix} m-1 & & -1 \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ -1 & & m-1 \end{bmatrix}, \tag{4.22}$$

and if we use the fact that

$$|Q| = \begin{vmatrix} a-1 & & -1 \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ -1 & & a-1 \end{vmatrix} = a^{k-1}[a-k],$$

where Q is a $k \times k$ matrix, it is thus seen that $\lambda_1 \leqq \theta_1$, which proves the theorem.

Acknowledgement. I would like to thank Professor Henry Scheffé for many helpful suggestions.

REFERENCES

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

[2] O. SCHREIER AND E. SPERNER, *Introduction to Modern Algebra and Matrix Theory*, Chelsea Publishing Company, New York, 1951.

[3] A. M. MOOD, "On Hotelling's weighing problem," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 432–447.

[4] G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, *Inequalities*, Cambridge, 1934, Section 2.13.

[5] C. R. RAO, "On the linear combination of observations and the general theory of least squares," *Sankhyā*, Vol. 7 (1946), pp. 237–256.

[6] F. N. DAVID AND J. NEYMAN, "Extension of the Markoff theorem of least squares," *Stat. Res. Memoirs*, Vol. 2 (1938), pp. 105–116.