THE CONSTRUCTION AND EVALUATION OF SOME DESIGNS FOR THE ESTIMATION OF PARAMETERS IN RANDOM MODELS

by

DAVID W. GAYLOR

and

R.L. ANDERSON

INSTITUTE OF STATISTICS

MIMEOGRAPH SERIES NO. 256

APRIL, 1960

Errata

p. 3, line 2 below (2): change "$n = \sum n_i$" to "$N = \sum n_i$".

p. 11, line 2 from botton: change "the" to "an".

p. 12, last line of Section 6.2: change "the" to "an".

p. 20, last line: change "b(blocks)" to "b blocks (columns)".

p. 65, Table 10: for E(MS) of Columns, change "$nc\sigma_c^2$" to "$nr\sigma_c^2$".

p. 80, line below (143): change "(172)" to "(143)".

p. 87, line 2: change "different" to "difficult".

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

1.0 LIST OF TABLES

3.0 INTRODUCTION

Considerable work has been done in the area of experimental design to estimate treatment contrasts. Estimates of components of variance can be obtained as a by-product of these designs. These estimates often may be equally or more important than estimates of the treatment contrasts. However, designs used primarily for the estimation of the latter may give very inefficient estimates of components of variance. Little work has been done in the general area of designing experiments especially for estimating components of variance. Anderson and Bancroft [1952, p. 334] introduce a "staggered" design as a possibility for use with nested classifications. Crump [1954] has considered optimal designs for estimating the components of variance for a one-way nested classification. In particular, Crump considered designs to estimate the component of variation among classes, σ_a^2, and the ratio $\rho = \sigma_a^2/\sigma_e^2$, where σ_e^2 is the within class component of variance. The intra-class correlation coefficient is $\sigma_a^2/(\sigma_a^2 + \sigma_e^2) = \rho/(1 + \rho)$.

The purpose of this thesis is to investigate designs for estimating the components of variance for a two-way crossed classification

$$y_{ijk} = \mu + r_i + c_j + (rc)_{ij} + e_{ijk} \qquad (1)$$

where $i = 1,2,\ldots r$; $j = 1,2,\ldots c$; $k = 1,2,\ldots n_{ij}$; and $\sum n_{ij} = N$, the total number of observations, is fixed. The general mean is μ; the row effects, r_i, are NID(0, σ_r^2); the column effects, c_j, are NID(0, σ_c^2); the interaction effects, $(rc)_{ij}$, are NID(0, σ_{rc}^2); the within cell effects, e_{ijk}, are NID(0, σ_e^2); and all effects are uncorrelated.

The optimal design for a particular problem will depend upon the method of estimation used. This thesis will emphasize the design of experiments for estimating components of variance rather than the method of estimation. In some cases, it will be possible to couple the best design with the best estimator. In other cases, the best estimator may be unknown. For these situations, a particular method of estimation will be selected and the allocation will be considered for that estimator.

The simultaneous estimation of variance components and sampling costs will also be investigated.

4.0 REVIEW OF LITERATURE

Hammersley [1949] and Crump [1954] considered optimal designs to estimate the parameters of the variance component model for a one-way nested classification

$$y_{ij} = \mu + a_i + e_{ij} \qquad (2)$$

where $i = 1,2,...a$; $j = 1,2,...n_i$; and the total number of observations, $N = \sum n_i$, is fixed. The class effects, a_i, are $NID(0,\sigma_a^2)$; the within class effects, e_{ij}, are $NID(0,\sigma_e^2)$; and all effects are uncorrelated. The analysis of variance for (2) is given in Table 1.

Table 1. Analysis of variance for a one-way nested classification [a]

Source of variation	DF	MS	E(MS)
Among classes	a-1	A	$\sigma_e^2 + n_o\sigma_a^2$
Within classes	N-a	W	σ_e^2
Total	N-1		

[a] DF is the degrees of freedom, MS is the mean square, and E(MS) is the expected value of the mean square.

Hammersley and Crump used the analysis of variance estimate of σ_a^2 obtained by equating the mean squares to their expected values in Table 1 giving

$$\hat{\sigma}_a^2 = (A-W)/n_o \ . \qquad (3)$$

Hammersley showed for a fixed a that the minimum variance of $\hat{\sigma}_a^2$ is obtained with equal numbers of observations per class. That is, N divided by a must give an integer, $n_i = n = N/a$. In this case, $n_o = n$, and the value of a which minimizes the variance of $\hat{\sigma}_a^2$ is

$$a_1 = N(N\rho + 2)/(N\rho + N + 1), \tag{4}$$

where $\rho = \sigma_a^2/\sigma_e^2$, with N, a, and n all integers. For this case with $n_i = n$ for all i, (3) is the maximum likelihood estimator and Graybill and Wortham [1956] show that (3) falls in the class of uniformly best (minimum variance) unbiased estimators. Hammersley offered no exact proofs of procedures to use for non-integers and only suggested using the nearest integers.

Crump extended these results to the situation where N/a is not necessarily an integer, but

$$N/a = p + s/a \tag{5}$$

where N, a, p, and s are integers with $o < s \leq a$. For fixed a, Crump showed that the variance of $\hat{\sigma}_a^2$ is minimized by setting p + 1 observations in s classes and p observations in the remaining a-s classes. Crump also shows that the variance of $\hat{\sigma}_a^2$ is minimized when the number of classes is given by (4). Crump hypothesized the following procedure for allocating samples to classes with fixed N:

(i) Find a_1 from (4). If a_1 is an integer, set $a = a_1$.

(ii) If a_1 is not an integer, let a be both the integer above and below a_1 and allocate for each value of a according to (5). Select the allocation which minimizes the variance of $\hat{\sigma}_a^2$. Hence, s classes would contain p + 1 observations and a-s classes p observations where p is the largest integer less than N/a. (Since the variance of $\hat{\sigma}_a^2$ differs very little for the two allocations, a general operational rule would be to choose a as that integer closest to a_1.)

Crump did not give a proof that this procedure always gives the optimal design, but for numerous examples this procedure always minimized the variance of $\hat{\sigma}_a^2$.

Baines [1944] used the F ratio, **A/W**, in Table 1 to obtain an estimator of $\rho = \sigma_a^2/\sigma_e^2$.

$$\tilde{\rho} = (\frac{A}{W} - 1)/n_o. \qquad (6)$$

Baines restricted his investigation to the case of equal numbers of observations per class, $n_i = n = n_o$, and found the value of n which minimized the variance of $\tilde{\rho}$. Crump considered the unbiased estimator of ρ, $\hat{\rho}$, which corrects the slight bias in $\tilde{\rho}$. Crump extended the results to the situation where N/a is a non-integer of the form given by (5). For a fixed a, Crump showed that the variance of $\hat{\rho}$ is minimized by putting p + 1 observations in s classes and p observations in a-s classes. The variance of $\hat{\rho}$ is minimized when the number of cells is

$$a' = 1 + \frac{(N - 5)(N\rho + 1)}{2N\rho + N-3} . \qquad (7)$$

It was hypothesized that the optimal design is obtained by setting a equal to the integer below or above a' which minimizes the variance of $\hat{\rho}$. Then, s classes contain p + 1 observations and a-s classes contain p observations where p is the largest integer less than N/a.

Crump shows that the guess or previous estimate of ρ required to determine a_1 in (4) and a' in (7) is not too critical. The design may differ considerably from the optimal design without materially affecting the variances of $\hat{\sigma}_a^2$ or $\hat{\rho}$.

Crump also shows that the optimal design for estimating μ is to set $n_i = 1$ and $a = N$ where the estimator is

$$\hat{\mu} = \sum y_{ij}/N.$$

The optimal design for estimating σ_e^2 is to set $n_i = N$ and $a = 1$ where the estimator is

$$\hat{\sigma}_e^2 = \sum (y_{ij} - \hat{\mu})^2/(N-1).$$

5.0 LOWER BOUND FOR THE VARIANCE OF UNBIASED ESTIMATES OF LINEAR COMBINATIONS OF COMPONENTS OF VARIANCE

Consider the general variance components model

$$y_i = \mu + e_i \tag{8}$$

where i = 1,2,...N; μ is the expected value of y_i; and the errors e_i are normally distributed with zero means and variance-covariance matrix V(NxN). Writing (8) in vector notation

$$\underline{y}\ (N\text{x}1) = \underline{\mu}\ (N\text{x}1) + \underline{e}\ (N\text{x}1) \tag{9}$$

where $E(\underline{e}) = \underline{0}$ (Nx1) and $E(\underline{ee}') = V$.

Consider a quadratic estimator

$$Q = \underline{y}'\ M\underline{y} \tag{10}$$

of a linear combination of components of variance, σ^2, where $M = M'$ (NxN).

Substituting from (9) into (10) gives

$$Q = (\underline{\mu}' + \underline{e}')\ M\ (\underline{\mu} + \underline{e}\ )$$

or

$$Q = \underline{\mu}'M\underline{\mu} + 2\underline{e}'M\underline{\mu} + \underline{e}'M\underline{e}. \tag{11}$$

Since $E(\underline{e}) = \underline{0}$, the expected value of Q is

$$E(Q) = \underline{\mu}'M\underline{\mu} + E(\underline{e}'M\underline{e}). \tag{12}$$

From Whittle [1953], Lancaster [1954], and others the s^{th} cumulant of $\underline{e}'M\underline{e}$ is

$$K_s = 2^{s-1}(s-1)!\ \text{tr}(VM)^s$$

or (13)

$$K_s = 2^{s-1}(s-1)! \sum \lambda_j^s$$

where λ (NxN) is the diagonal matrix of latent roots of VM. From (13)

$$E(\underline{e}'M\underline{e}) = tr(VM) ,$$

giving from (12)

$$E(Q) = \underline{\mu}'M\underline{\mu} + tr(VM). \tag{14}$$

Squaring (11) gives

$$Q^2 = (\underline{\mu}'M\underline{\mu})^2 + 4(\underline{e}'M\underline{\mu})^2 + (\underline{e}'M\underline{e})^2 + 4(\underline{\mu}'M\underline{\mu})(\underline{e}'M\underline{\mu}) + 2(\underline{\mu}'M\underline{\mu})(\underline{e}'M\underline{e}) + 4(\underline{e}'M\underline{\mu})(\underline{e}'M\underline{e}). \tag{15}$$

Since the expected values of odd powers of $\underline{e}$ vanish for the normal distribution

$$E(Q^2) = (\underline{\mu}'M\underline{\mu})^2 + 4\underline{\mu}'MVM\underline{\mu} + E(\underline{e}'M\underline{e})^2 + 2\underline{\mu}'M\underline{\mu}\ E(\underline{e}'Me). \tag{16}$$

The variance of Q is

$$V(Q) = E(Q^2) - [E(Q)]^2.$$

From (12) and (16)

$$V(Q) = 4\underline{\mu}'MVM\underline{\mu} + E(\underline{e}'M\underline{e})^2 - [E(\underline{e}'M\underline{e})]^2;$$

however,

$$E(\underline{e}'M\underline{e})^2 - [E(\underline{e}'M\underline{e})]^2$$

is the second cumulant of $\underline{e}'M\underline{e}$, which from (13) is $2\ tr(VM)^2$. Thus,

$$V(Q) = 4\underline{\mu}'MVM\underline{\mu} + 2\ tr(VM)^2. \tag{17}$$

Restricting Q to unbiased estimators requires

$$E(Q) = \underline{\mu}'M\underline{\mu} + tr(VM) = \sigma^2 . \tag{18}$$

V and M do not contain $\underline{\mu}$ and M does not contain any estimates from the data of the variances or covariances of the e_i's. Thus, $\underline{\mu}'M = \underline{0}'$. That is, each row and column of M must sum to zero. Thus,

$$\text{rank } (M) = q \leq N-1.$$

Since $\underline{\mu}'M = \underline{0}'$, (17) and (18) become

$$E(Q) = tr(VM) = \sigma^2 \tag{19}$$

and

$$V(Q) = 2\ tr(VM)^2. \tag{20}$$

Since $tr(VM)^s = tr(\lambda)^s$, where $\lambda(N\text{x}N)$ is the diagonal matrix of the latent roots of VM with rank = q,

$$tr(\lambda) = \sum \lambda_j = \sigma^2 \tag{21}$$

and

$$V(Q) = 2\ tr(\lambda)^2 = 2\sum \lambda_j^2. \tag{22}$$

It is desired to minimize V(Q) subject to the restriction that $E(Q) = tr(\lambda) = \sigma^2$. Let

$$P = 2\sum \lambda_j^2 - L\ (\sum \lambda_j - \sigma^2)$$

where L is the LaGrange multiplier. Setting the partial derivative of P with respect to the j^{th} non-zero latent root of VM equal to zero gives

$$\frac{\delta P}{\delta \lambda_j} = 4\ \lambda_j - L = 0;$$

hence,

$$\lambda_j = L/4. \tag{23}$$

Summing (23) over j,

$$\sum_{j=1}^{q} \lambda_j = \sum \frac{L}{4} = \frac{qL}{4}, \tag{24}$$

since

$$\sum \lambda_j = tr(\lambda) = \sigma^2,$$

$$qL/4 = \sigma^2$$

or

$$L = 4\ \sigma^2/q\ .$$

Substituting this result into (23),

$$\lambda_j = \sigma^2/q \, , \qquad j = 1,2,\ldots q \, ,$$

and from (22) the minimum variance of Q is then

$$\min V(Q) = 2 \, tr(\lambda)^2 = 2 \sum \lambda_j^2 = 2\sigma^4/q. \tag{25}$$

It was shown that the maximum value that q can have is N-1 for an unbiased estimator of σ^2. Thus, the lower bound for the variance of an unbiased estimator of a linear combination of components of variance, σ^2, is

$$L.B.V(Q) = 2 \, \sigma^4/(N-1) \, . \tag{26}$$

The lower bound usually will not be attainable. However, if a design is found for which an unbiased estimator has a variance of $2 \, \sigma^4/(N-1)$, then it is impossible to find an unbiased estimator and design which give a lower variance.

6.0 OPTIMAL DESIGNS FOR ESTIMATING COMPONENTS OF VARIANCE IN A TWO-WAY CROSSED CLASSIFICATION

The first general type of problem considered is the optimal design for estimating a single function of components of variance in the two-way crossed classification given by model (1). The functions considered are single components of variance, sums of components, and ratios of components. In this section it will be assumed that it is no more costly to sample different rows or columns than to sample within a cell. That is, the cost of sampling is directly proportional to N, the total number of observations. Optimal designs for simultaneously estimating two components of variance are considered in later sections.

6.1 Estimation of μ

Consider the following unbiased estimator of the general mean for the model given by (1)

$$\hat{\mu} = \sum_{i,j}^{N} y_{ij}/N.$$

The variance of $\hat{\mu}$ is

$$v(\hat{\mu}) = \frac{\sigma_e^2}{N} + \frac{\sigma_{rc}^2 \sum\sum n_{ij}^2}{N^2} + \frac{\sigma_r^2 \sum n_{i\cdot}^2}{N^2} + \frac{\sigma_c^2 \sum n_{\cdot j}^2}{N^2} , \qquad (27)$$

where $n_{i\cdot} = \sum_j n_{ij}$ and $n_{\cdot j} = \sum_i n_{ij}$. The values of $\sum\sum n_{ij}^2$, $\sum n_{i\cdot}^2$, and $\sum n_{\cdot j}^2$ are minimized when $n_{ij} = 0$ or 1 and $n_{i\cdot} = n_{\cdot j} = 1$. That is, ~~the~~ an optimal design is to select each observation from a different row and column, giving

$$V(\hat{\mu}) = (\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2 + \sigma_c^2)/N ,$$

which is best since each component of variance is divided by N, the maximum possible number of observations.

6.2 Estimation of σ_e^2

Consider the following unbiased estimator of σ_e^2 for the model given by (1) with N observations in one cell

$$\hat{\sigma}_e^2 = \sum_{k=1}^{N} (y_{ijk} - y_{ij\cdot})^2/(N-1) ,$$

where $y_{ij\cdot} = \sum_{k=1}^{N} y_{ijk}/N$, the mean of the ij cell. The variance of $\hat{\sigma}_e^2$ is

$$V(\hat{\sigma}_e^2) = 2\ \sigma_e^4/(N-1) .$$

Since this attains the lower bound for the variance given by (26), an optimal design is to place all N observations in a single cell.

6.3 Estimation of $(\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2)$ or $(\sigma_e^2 + \sigma_{rc}^2 + \sigma_c^2)$

Consider the following unbiased estimator of $(\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2)$ with observations in one column and N rows

$$(\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2) \hat{=} \sum_{i=1}^{N} (y_{ij} - y_{\cdot j})^2/(N-1)$$

where $y_{\cdot j} = \sum_{i=1}^{N} y_{ij}/N$, the mean of the j^{th} column and $\hat{=}$ means "is estimated by". The variance of this estimator is $2\ (\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2)^2/(N-1)$.

Since this variance attains the lower bound given by (26), the optimal design is to select observations from only one column and N different rows in that column.

Similarly, the best design for estimating $(\sigma_e^2 + \sigma_{rc}^2 + \sigma_c^2)$ is obtained by sampling from only one row and N different columns in that row.

6.4 Estimation of $(\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2 + \sigma_c^2)$

The total variance of an observation, y_{ijk}, from (1) is $(\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2 + \sigma_c^2)$. Consider the following unbiased estimator of the total variance with each observation from a different row and column

$$(\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2 + \sigma_c^2) \hat{=} \sum_{i=1}^{N} (y_{ii} - y_{..})^2/(N-1) ,$$

where $y_{..} = \sum_{i=1}^{N} y_{ii}/N$. The variance of this estimator is

$$2\ (\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2 + \sigma_c^2)^2/(N-1).$$

Since this variance attains the lower bound for the variance of an estimator given by (26), the optimal design is to select each observation from a different row and column.

6.5 Estimation of σ_r^2 or σ_c^2

For estimating either σ_r^2 or σ_c^2, we have not been able to develop a general class of designs which can be proven to be optimal for all situations. Instead the discussion has been limited to optimal designs for the class of connected designs with 0 or n observations per cell, i.e. $n_{ij} = 0$ or n. A connected design is one for which the adjusted sum of squares for rows, as presented in Table 2, has r-1 degrees of freedom. Similarly, when estimating σ_c^2, the adjusted sum of squares for columns has c-1 degrees of freedom. The analysis of variance, which is obtained by the method of fitting constants, is given in Table 2 for the model given by (1).

Table 2. Analysis of variance for a two-way crossed classification with $n_{ij} = 0$ or n

Source of variation	DF	MS	E(MS)
Columns	c-1	C	$\sigma_e^2 + n\sigma_{rc}^2 + c_1\sigma_r^2 + r_o\sigma_c^2$
Rows (adjusted for columns)	r-1	R*	$\sigma_e^2 + n\sigma_{rc}^2 + c_o\sigma_r^2$
Interaction (adjusted for rows and columns)	$\frac{N}{n}$-r-c+1	I*	$\sigma_e^2 + n\sigma_{rc}^2$
Error	$N-\frac{N}{n}$	E	σ_e^2
Total	N-1		

An unbiased estimator of σ_r^2 can be obtained by equating the mean squares to their expected values in Table 2 giving

$$\hat{\sigma}_r^2 = (R^* - I^*)/c_o \quad , \tag{28}$$

where the value of c_o in general is given by Kempthorne [1952, p. 112]

$$c_o = (N - \sum_{j}^{c}\sum_{i}^{r} n_{ij}^2/n._j)/(r-1) \quad , \qquad r \geq 2$$

where

$$n._j = \sum_{i}^{r} n_{ij} \quad .$$

For the case being considered here with $n_{ij} = 0$ or n, c_o is

$$c_o = (N-nc)/(r-1) \quad . \tag{29}$$

Alternative analyses of variance would lead to alternative estimators of σ_r^2. The investigation here will be limited to the estimator given by (28). It is realized that when the n_{ij} are not all equal, other estimators may be more efficient than the one proposed here. It is hoped to study this problem in subsequent investigations. However, for this thesis, we will consider only (28) because this is one estimator in common usage and it has optimal properties when all n_{ij} are equal.

In Section 5.0, a general quadratic estimator, Q, was considered for estimating a linear function of components of variance,

$$Q = \underline{y}'M\underline{y}$$

expressed in matrix notation. By an orthogonal transformation, Q can be expressed as

$$Q = \sum \lambda_i \chi_i^2 \tag{30}$$

where the χ_i^2 are independent chi-square variates each with one degree of freedom, and the λ_i's are the latent roots of VM, where V is the variance-covariance matrix of the y's. Also, in Section 5.0, it was shown that the expected value of Q is

$$E(Q) = \sum \lambda_i$$

and the variance of Q is

$$V(Q) = 2\sum \lambda_i^2 \quad . \tag{31}$$

In general, the row sum of squares adjusted for columns, $(r-1)R^*$, can be divided into its orthogonal individual sums of squares

$$(r-1)R^* = \sum_{j=1}^{r-1} (\sigma_e^2 + n\sigma_{rc}^2 + nc_j\sigma_r^2)\chi_j^2 \quad , \tag{32}$$

where the χ_j^2 are independent chi-square variates, each with one degree of freedom. Similarly,

$$(N/n - r-c+1)I = \sum_{k=1}^{N/n - r-c+1} (\sigma_e^2 + n\sigma_{rc}^2)\chi_k^2 \quad . \tag{33}$$

From (28), $Q = \hat{\sigma}_r^2$, which may be expressed in general as

$$\hat{\sigma}_r^2 = \sum_{j=1}^{r-1} \lambda_j\chi_j^2 + \sum_{k=1}^{N/n - r-c+1} \lambda_k\chi_k^2 \quad , \tag{34}$$

where

$$\lambda_j = \frac{\sigma_e^2 + n\sigma_{rc}^2 + nc_j\sigma_r^2}{N-nc}$$

and

$$\lambda_k = -\frac{(r-1)(\sigma_e^2 + n\sigma_{rc}^2)}{(N-nc)(N/n - r-c+1)} \quad .$$

Since

$$E(\hat{\sigma}_r^2) = \sigma_r^2 \quad ,$$

$$\sum \lambda_j + \sum \lambda_k = \sigma_r^2 \quad ;$$

hence,

$$\sum_{j=1}^{r-1} c_j = (N-nc)/n = c_o(r-1)/n \quad . \tag{35}$$

From (31) and (34), the variance of $\hat{\sigma}_r^2$ is

$$V_n(\hat{\sigma}_r^2) = 2\sum_{j=1}^{r-1} \frac{(\sigma_e^2 + n\sigma_{rc}^2 + nc_j\sigma_r^2)^2}{(N-nc)^2} + \frac{2(r-1)^2\ (\sigma_e^2 + n\sigma_{rc}^2)^2}{(N-nc)^2(N/n - r-c+1)} \quad . \tag{36}$$

It will now be shown that a design with $n_{ij} = 0$ or 1 can be constructed which has a variance of $\hat{\sigma}_r^2$ smaller than or equal to the variance of $\hat{\sigma}_r^2$ from a design with $n_{ij} = 0$ or n. Consider any two-way design with r rows, c columns, $n_{ij} = 0$ or n, and $\sum n_{ij} = N$. Expand the j^{th} column into n identically filled columns with zero observations in each of the n columns for the i^{th} row if $n_{ij} = 0$ and one observation in each of the n columns for the i^{th} row if $n_{ij} = n$. Each column is expanded in this manner for all $j = 1,2,\ldots c$. In this way, a design with $n_{ij}^* = 0$ or 1 is constructed which has r rows, $c^* = nc$ columns, and $\sum n_{ij}^* = N$. It follows from (36) that the variance of $\hat{\sigma}_r^2$ for this design is

$$V^*(\hat{\sigma}_r^2) = 2\sum_{j=1}^{r-1} \frac{(\sigma_e^2 + \sigma_{rc}^2 + c_j^*\sigma_r^2)^2}{(N-c^*)^2} + \frac{2(r-1)^2(\sigma_e^2 + \sigma_{rc}^2)^2}{(N-c^*)^2(N - r-c^*+1)} \quad , \tag{37}$$

where $c^* = nc$ and $c_j^* = nc_j$. Since

$$N-c^* = N-nc \geq (N-nc)/n \quad ,$$

$$N-c^*-r+1 \geq N/n - c-r+1 \quad . \tag{38}$$

Using the relationships $c^* = nc$ and $c_j^* = nc$; and (38), the variance of $\hat{\sigma}_r^2$ given by (37) with $n_{ij} = 0$ or 1 is always less than or equal to the variance of $\hat{\sigma}_r^2$ given by (36) where $n_{ij} = 0$ or n. That is, from any design with $n_{ij} = 0$ or n, a design with $n_{ij} = 0$ or 1 can be constructed which has an equivalent or smaller variance for $\hat{\sigma}_r^2$.

The problem now remains of finding the design with $n_{ij} = 0$ or 1 which minimizes the variance of $\hat{\sigma}_r^2$.

Choose any $c'(2 \leq c' \leq N-r)$, the maximum number of columns sampled. Let c_j' be the coefficients in the orthogonal forms corresponding to (32) for n=1. Similarly,

$$\lambda_j' = \frac{\sigma_e^2 + \sigma_{rc}^2 + c_j'\sigma_r^2}{N-c'} , \qquad j = 1,2,\ldots, r-1;$$

$$\lambda_k' = -\frac{(r-1)(\sigma_e^2 + \sigma_{rc}^2)}{(N-c')(N-r-c'+1)} , \qquad k = 1,2,\ldots, N-r-c'+1.$$

Hence,

$$\sum_1^{r-1} c_j' = N-c' = c_o'(r-1) \quad . \tag{39}$$

The variance of $\hat{\sigma}_r^2$ may be written in the following form:

$$V'(\hat{\sigma}_r^2) = \frac{2(r-1)\sigma^4}{(N-c')(N-r-c'+1)} + \frac{4\sigma^2\sigma_r^2}{N-c'}$$

$$+ \frac{2\sigma_r^4}{(N-c')^2} \sum_{j=1}^{r-1} c'^2_j \quad ,$$

where $\sigma^2 = \sigma_e^2 + \sigma_{rc}^2$. This expression is obviously minimized when both c' and $\sum c'^2_j$ are small.

For a given N, r, and c', we wish to allocate the samples to the given rows and columns so as to minimize $\sum c_j'^2$, subject to the restriction that $\sum c_j' = N - c' = c_o'(r-1)$. We note that

$$\sum c_j'^2 = \sum c_j'^2 - \frac{(N-c')^2}{r-1} + \frac{(N-c')^2}{r-1}$$

$$= \sum (c_j' - c_o')^2 + (N-c')^2/(r-1) \quad .$$

Hence, $V'(\hat{\sigma}_r^2)$ becomes

$$V'(\hat{\sigma}_r^2) = \frac{2(r-1)\sigma^4}{(N-c')(N-r-c'+1)} + \frac{4\sigma^2\sigma_r^2}{N-c'}$$
$$+ \frac{2\sigma_r^4}{(N-c')^2} \sum_{j=1}^{r-1} (c_j' - c_o')^2 + \frac{2\sigma_r^4}{r-1} \quad . \tag{40}$$

Equation (40) is minimized, for a given N, r and c', if the variance of c_j' is as small as possible. This is accomplished by making the c_j''s as nearly alike as possible. If N is divisible by r, such that $N/r = c$, an integer, then $c' = c = c_o'$ is the minimum number of columns and $\sum (c_j' - c_o')^2 = 0$. Therefore, for a given r, the variance of $\hat{\sigma}_r^2$ as given by (40) is minimized. In this case, the optimal design consists of c columns each with r rows in common.

If N/r is not an integer, we write

$$N = r(c-1) + s \quad , \qquad 0 < s \leq r \quad . \tag{41}$$

In this case, consider the design shown in Figure 1 which has the property of having the smallest value of $c'(c' = c + s - 1)$ for which $\sum (c_j' - c_o')^2$ is zero. For the design shown in Figure 1, the c_j' $(j = 1,2,\ldots, r-1)$ are identified by (39) with the numbers of columns per row, c-1, in the first r-1 rows, i.e., $c_j' = (N-c')/(r-1)$ $= \left[r(c-1) + s - (c+s-1)\right]/(r-1) = c-1$.

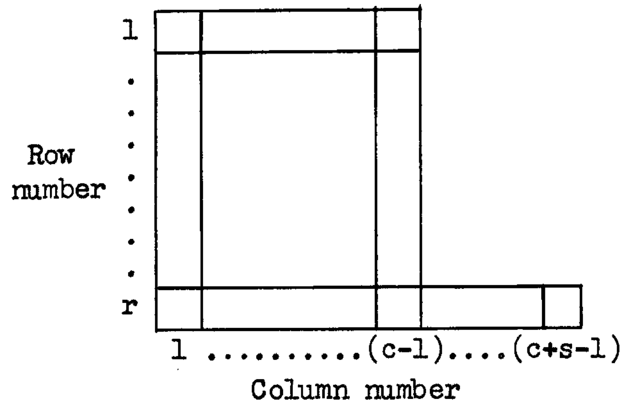

Figure 1. Design for $s \leq r$, $c' = c+s-1$

In Figure 1, if any observations are taken from the first c-1 columns (subject to the condition that r and c' are not decreased) and added to the columns to the right of the c-1 columns in Figure 1, $\sum (c_j' - c_o')^2 \geq 0$ and $c' \geq c+s-1$. Hence, the variance of $\hat{\sigma}_r^2$ as seen by (40) would be greater than or equal to the variance of $\hat{\sigma}_r^2$ for the configuration shown in Figure 1.

The only designs remaining which may be better than Figure 1 are those obtained by taking some of the observations in the last s columns and placing them in fewer columns. As the number of columns is decreased in this manner, N-c' is increased, which decreases the variance of $\hat{\sigma}_r^2$. But, $\sum (c_j' - c_o')^2$ becomes greater than zero which increases the variance. The solution for c' depends on the relative sizes of σ_r^2 and σ^2. This final question of allocating these s observations to minimize the variance of $\hat{\sigma}_r^2$ will now be investigated.

A result obtained by Shah [1959] will be used. Shah considered connected balanced incomplete designs with v treatments (rows) each appearing r times and b (blocks) (columns) of k plots each. Treatment i appears n_{ij} times in block j (n_{ij} = 0 or 1). In this case $A\hat{t} = Q$, where $Q = T - (1/k)NB$, $A = r\,I(vxv) - (1/k)NN'$, $N = (n_{ij})$, and T and B represent vectors of treatment and block totals. Shah shows that

$$\hat{t} = \left[A + a\,E(vxv)\right]^{-1} Q$$

and

$$A\left[A + a\,E(vxv)\right]^{-1} = I(vxv) - (1/v)\,E(vxv) \quad , \qquad (42)$$

where a is any non-zero real number and E(vxv) is a matrix of one's. Apparently, these results extend to any incomplete blocks situation regardless of the values of $n_{i.}$ and $n_{.j}$ where $n_{i.} = \sum_j n_{ij}$ and $n_{.j} = \sum_i n_{ij}$. For our problem

$$A(rxr)\,R(rx1) = Q(rx1) \quad , \qquad (43)$$

where R is a vector of $NID(0,\sigma_r^2)$ variates, A is the adjusted sum of squares and products matrix for rows, and Q is the adjusted vector of row totals (all adjusted for columns). In the least squares sense

$$\hat{R} = \left[A + a\,E(rxr)\right]^{-1} Q \quad ;$$

since (A + aE) is symmetrical,

$$SSR^* = Q'\,(A + aE)^{-1}\,Q \quad .$$

Substituting from (43),

$$SSR^* = R'A\,(A + aE)^{-1}\,AR \quad .$$

From (42) this becomes

$$SSR^* = R'AR - (1/v)\,R'EAR \quad .$$

Since the sum of each column of the A matrix is zero,

$$EA = 0 \quad .$$

Thus,

$$SSR^* = R'AR = \sum_{i=1}^{r-1} \lambda_i \chi_i^2 \quad ,$$

where the χ_i^2 are independent chi-square variates with one degree of freedom each.

For the design being considered, the rows can be permuted so that the first r-t rows have c-1 columns in common and the last t rows have the same c-1 columns in common plus s observations distributed in columns, c, c+1,..., c'-1, c'. In this case the matrix of sums of squares and cross products in the normal equations is

	$c_1 \ldots c_{c-1}$	$c_c \ldots c_{c'}$	$r_1 \ldots r_{r-t}$	$r_{r-t+1} \ldots r_r$
$c_1 \ldots c_{c-1}$	$r \ldots r$ (diagonal), 0	0	$1 \ldots 1$ (all 1)	$1 \ldots 1$ (all 1)
$c_c \ldots c_{c'}$		$n_{\cdot c} \ldots n_{\cdot c'}$ (diagonal), 0	0	$n_{c,r-t+1} \ldots n_{c,r}$; $\ldots$; $n_{c',r-t+1} \ldots n_{c',r}$
$r_1 \ldots r_{r-t}$			$c-1 \ldots c-1$ (diagonal), 0	0
$r_{r-t+1} \ldots r_r$				$n_{1\cdot}+c-1 \ldots n_{t\cdot}+c-1$ (diagonal), 0

where the transpose of these elements appear below the diagonal. The A matrix of the adjusted row sum of squares and cross products is

$$A(r\times r) = \left[\begin{array}{c|c} A_{11}\ (r-t)\times(r-t) & A_{12}\ (r-t)\times t \\ \hline A_{12}'\ t\times(r-t) & A_{22}\ (t\times t) \end{array}\right].$$

The sub-matrices are

$$A_{11} = (c-1)\ I - \frac{(c-1)}{r}\ E$$

$$A_{12} = -\frac{(c-1)}{r}\ E$$

$$A_{22} = (c-1)\ I - \frac{(c-1)}{r}\ E + B$$

where

$$B = \begin{bmatrix} \left(n_{1\cdot} - \sum_i \frac{n_{i,\,r-t+1}^2}{n_{\cdot i}}\right) & \cdots & \left(-\sum_i \frac{n_{i,\,r-t+1}\, n_{ir}}{n_{\cdot i}}\right) \\ \vdots & \ddots & \vdots \\ \left(-\sum_i \frac{n_{i,\,r-t+1}\, n_{ir}}{n_{\cdot i}}\right) & \cdots & \left(n_{t\cdot} - \sum_i \frac{n_{ir}^2}{n_{\cdot i}}\right) \end{bmatrix}$$

Adding (c-1)/r to every element in A gives

$$A + \frac{(c-1)}{r}\ E = \left[\begin{array}{c|c} (c-1)\ I & 0 \\ \hline 0 & (c-1)\ I + B \end{array}\right].$$

Shah shows that by using this technique, the latent roots of $[A + aE(r\times r)]$ are also the latent roots of A except that the zero root of A becomes ar. The above matrix has r-t roots of c-1 from the first r-t rows. Let

$$D = (c-1)\ I + B$$

then

$$\left| D - (c-1)\, I \right| = \left| B \right| .$$

Since $\left| B \right| = 0$,

$$\left| D - (c-1)\, I \right| = 0 \quad .$$

Hence, at least one root from the last t rows is c-1. Therefore, $(A + \frac{c-1}{r} E)$ has at least r-t+1 roots of c-1. One of these roots resulted from the technique of adding (c-1)/r. Using the results of Shah, A has at least r-t latent roots of c-1. That is, $c_j' = c-1$ (j=1,2,...r-t).

Alternatively, from

$$\begin{aligned} \sum_{j=1}^{r-1} c_j' &= N - c' \\ &= r(c-1) + s - (c+s-t) \\ &= r(c-1) - c+t. \end{aligned} \qquad (44)$$

Similarly, from the last t rows

$$\begin{aligned} \sum_{r-t+1}^{r-1} c_j' &= t(c-1) + s - (c+s-t) \\ &= c(t-1) \quad . \end{aligned}$$

Due to the balance from the first r-t rows

$$c_j' = c-1 \qquad\qquad j = 1,2,\ldots., r-t-1.$$

The remaining root is

$$\begin{aligned} c'_{r-t} &= \sum_{j=1}^{r-1} c_j' - \sum_{j=1}^{r-t-1} c_j' - \sum_{r-t+1}^{r-1} c_j' \\ &= r(c-1) - c + t - (r-t-1)(c-1) - c(t-1) \\ &= c-1. \end{aligned}$$

Hence, there are at least r-t values of c_j' equal to c-1.

Consider the design shown in Figure 2 which is constructed from the design in Figure 1 by placing t-1 ($1 \le t \le s$) observations from the last t-1 columns into the c^{th} column. Then, $c' = c + s - t$.

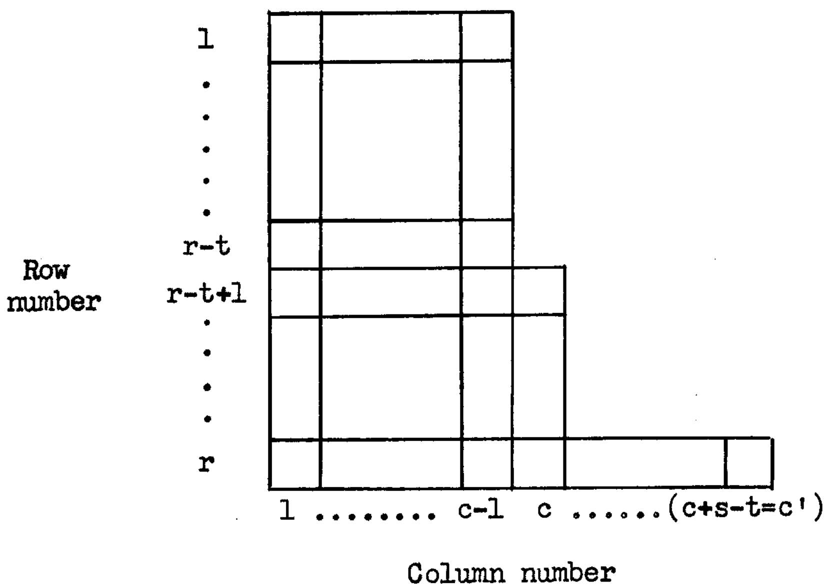

Figure 2. Design for $1 \le t \le s \le r$, $c' = c+s-t$

It has been shown previously that this design has at least r-t values of c_j' equal to c-1. Also, t-2 values of c_j' equal to c are obviously obtained from the t-1 rows r-t+1, r-t+2, ..., r-2, r-1. From (44), the remaining value of c_j' is

$$c_{r-1}' = \sum_{1}^{r-1} c_j' - (r-t)(c-1) - (t-2)c = c \quad .$$

Hence,

$$c_j' = c-1 \qquad j = 1,2,\ldots, r-t$$

$$c_j' = c \qquad j = r-t+1,\ldots, r-1$$

for the design shown in Figure 2. For a fixed c', the variance of

$\hat{\sigma}_r^2$ is minimized by minimizing $\sum (c'_j - c'_o)^2$. It will now be shown that this is accomplished by the design shown in Figure 2. The corrected sum of squares among the c'_j's is the same as the corrected sum of squares among the coded values $c''_j = c'_j - (c-1)$. Then,

$$\sum_{j=1}^{r-1} (c'_j - c'_o)^2 = \sum_{j=1}^{r-1} c''^2_j - \frac{(t-1)^2}{r-1} \quad , \tag{45}$$

where $1 \leq t \leq s \leq r$. For any particular c' (or t), $\sum (c'_j - c'_o)^2$ is minimized by minimizing $\sum c''^2_j$. The sum of the c''_j's is

$$\sum_{j=1}^{r-1} c''_j = \sum_{j=1}^{r-1} c'_j - (r-1)(c-1) \quad .$$

From (39)

$$\begin{aligned} \sum_{j=1}^{r-1} c'_j &= N-c' \\ &= r(c-1) + s - (c+s-t) \\ &= r(c-1) - c + t \quad . \end{aligned} \tag{46}$$

Then $\sum c''_j$ becomes

$$\begin{aligned} \sum_{j=1}^{r-1} c''_j &= r(c-1) - c + t - (r-1)(c-1) \\ &= t - 1 \quad , \end{aligned} \tag{47}$$

and

$$\bar{c}'' = (t-1)/(r-1) \quad .$$

We note that

$$\sum_{j=1}^{r-1} (c''_j - \bar{c}'')^2 = \begin{bmatrix} \text{SS among first} \\ r-t,\ c''_j\text{'s} \end{bmatrix} + \begin{bmatrix} \text{SS among last} \\ t-1,\ c''_j\text{'s} \end{bmatrix} + \begin{bmatrix} \text{SS between first group} \\ \text{and second group} \end{bmatrix}$$

or

$$\sum_{j=1}^{r-1} (c''_j-\bar{c}'')^2 = \sum_{j=1}^{r-t} (c''_j-\bar{c}''_1)^2 + \sum_{j=r-t+1}^{r-1} (c''_j-\bar{c}''_2)^2 + \frac{\left[(r-t)\bar{c}''_1\right]^2}{r-t} + \frac{\left[(t-1)\bar{c}''_2\right]^2}{t-1} - \frac{\left[(r-1)\bar{c}''\right]^2}{r-1} \qquad (48)$$

where

$$\bar{c}''_1 = \sum_1^{r-t} c''_j/(r-t)$$

and

$$\bar{c}''_2 = \sum_{r-t+1}^{r-1} c''_j/(t-1) \quad .$$

Due to the balanced portions in Figure 2 the latent roots are easily identified giving

$$c''_j = 0 \qquad j = 1,2,\ldots, r-t$$

$$c''_j = 1 \qquad j = r-t+1,\ldots, r-1 \quad .$$

Then, in (48)

$$\sum_1^{r-t} (c''_j - \bar{c}''_1)^2 = 0$$

and

$$\sum_{r-t+1}^{r-1} (c''_j - \bar{c}''_2)^2 = 0 \quad .$$

Also

$$\sum_1^{r-t} c''_j = 0$$

and from (47)

$$\sum_{r-t+1}^{r-1} c''_j = t - 1 \quad .$$

Thus,

$$\begin{bmatrix}\text{SS between first and} \\ \text{second group of } c_j'''\text{s}\end{bmatrix} = (t-1) - \frac{(t-1)^2}{r-1} = \frac{(t-1)(r-t)}{r-1} \quad . \tag{49}$$

Substituting these results into (48), for the design shown in Figure 2

$$\sum_{1}^{r-1} (c_j'' - \bar{c}'')^2 = \frac{(t-1)(r-t)}{(r-1)} \quad . \tag{50}$$

Suppose the t-1 observations are allotted to any of the s-t+1 columns beyond the first c-1 columns. The rows can be permuted so that at least the first r-t rows have c-1 columns in common. Hence, there are always at least r-t values of $c_j' = c-1$ or $c_j'' = 0$. When the t-1 observations are allotted to any of the s-t+1 columns beyond the first c-1 columns

$$\sum_{1}^{r-t} (c_j'' - \bar{c}_1'')^2 = 0 \quad ,$$

$$\sum_{r-t+1}^{r-1} (c_j'' - \bar{c}_2'')^2 \geq 0 \quad ,$$

$$\begin{bmatrix}\text{SS between the two} \\ \text{groups of } c_j'''\text{s}\end{bmatrix} = \frac{(t-1)(r-t)}{r-1}$$

as in (50) since $\sum_{1}^{r-t} c_j'' = 0$ and from (47) $\sum_{r-t+1}^{r-1} c_j'' = t-1$ as before.

Substituting these results into (48) gives

$$\sum_{1}^{r-1} (c_j'' - \bar{c}'')^2 \geq \frac{(t-1)(r-t)}{r-1} \quad , \tag{51}$$

when the t-1 observations are allotted to different columns. From (50), the equality holds when the t-1 observations are all allotted to the same column. Hence, for any particular value of c', the quantity $\sum (c_j'' - \bar{c}'')^2$ for the allocation shown in Figure 2 is less than or equal to the quantity obtained by any other allocation of the t-1 observations to the last s-t+1 columns.

In Figure 2, consider for any $1 \leq t \leq s \leq r$ the observations in the last s-t columns of the r^{th} row. These observations are completely confounded with columns and contribute no information to the estimation of σ_r^2. Suppose these s-t observations are deleted from the design. As seen previously, $\sum (c_j' - c_o')^2$ will remain unaffected by the deletion since the c_j' are identified with the numbers of observations per row in the first r-1 rows. Also, if the s-t observations are deleted, then s-t columns are deleted so that $N-c' = [N - (s-t)] - [c' - (s-t)]$ is unaffected. Therefore, the variance of $\hat{\sigma}_r^2$ as given by (40) is not changed by deleting the last s-t columns in Figure 2. Thus, the optimal design for fixed r is of the form shown in Figure 3, consisting of c-1 columns with r rows in common and one column with u of the r rows, where $0 \leq u \leq s \leq r$ $(u \neq 1)$ and r, c, and s are defined by (41). That is,

$$u = t \text{ for } 2 \leq t \leq s \leq r$$

$$u = 0 \text{ for } t = 1$$

since in the latter case, s observations are discarded.

From (37), the variance of $\hat{\sigma}_r^2$ for the optimal design shown in Figure 3 is

$$V(\hat{\sigma}_r^2) = \frac{2}{(N-c')^2} \left[\begin{array}{l} (r-t)\left[\sigma^2 + (c-1)\sigma_r^2\right]^2 + (t-1)(\sigma^2 + c\sigma_r^2) \\ + (r-1)^2 \sigma^4/(N-r-c'+1) \end{array} \right] \tag{52}$$

where $t = u$ for $t \geq 2$ and $t = 1$ for $u = 0$, and the total sample size actually used is $N^* = N-s+u$.

Similarly, it can be shown that placing the r-u empty cells in more than one of the c columns increases $\sum (c_j' - c_o')^2$ and hence increases the variance of $\hat{\sigma}_r^2$.

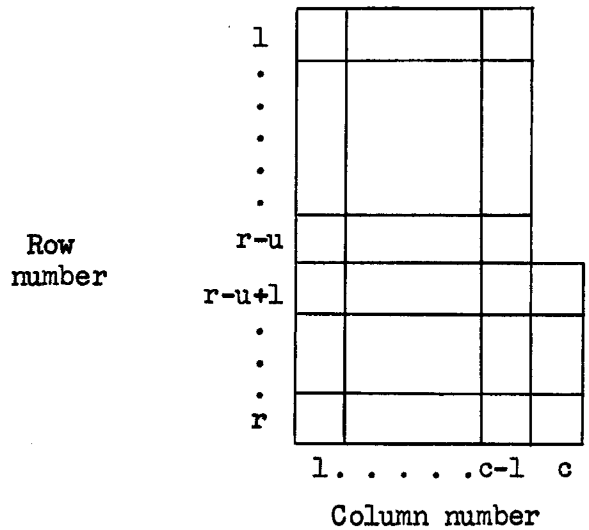

Figure 3. Optimal design, $0 \leq u \leq s \leq r$ $(u \neq 1)$

The question now remains of finding the best value of t for any particular N and r. As observations are taken from the last columns and placed in the c^{th} column, $N-c'$ increases, which decreases the variance of $\hat{\sigma}_r^2$. But, $\sum (c_j' - c_o')^2$ becomes greater than zero which increases the variance. For the form of the optimal design shown in Figure 3, substituting the result from (50) into (40) gives for the variance of $\hat{\sigma}_r^2$

$$V(\hat{\sigma}_r^2) = \frac{2\sigma^4}{(N-c')}\left[\frac{r-1}{(N-c'-r+1)} + 2\rho + \frac{\rho^2(r-t)(t-1)}{(N-c')(r-1)} + \frac{\rho^2(N-c')}{r-1}\right] \quad (53)$$

where $\rho = \sigma_r^2/\sigma^2$, $\sigma^2 = \sigma_e^2 + \sigma_{rc}^2$, $N = r(c-1) + s$, $1 \leq s \leq r$, $1 \leq t \leq s \leq r$, and $c' = c + s - t$. The term $(r-t)(t-1)$ is quadratic in t and reaches a maximum at $t = r/2 = (r+2)/2$ (r even) and $t = (r+1)/2$ (r odd). This term increases monotonically from $t = 1$ to $t = (r+1)/2$ and then decreases monotonically to $t = r$. The last term in (53) is constant for all t and the first two terms of (53) are monotonically decreasing as t increases. Thus, the variance of $\hat{\sigma}_r^2$ is monotonically decreasing for $t \geq (r+1)/2$.

Since

$$N - c' = r(c-1) + t \quad ,$$

the best design when $s = r$ is to set $t = s = r$. This clearly minimizes the variance as given by (53), and it uses all of the N possible samples. Also, in this balanced case, Graybill and Wortham [1956] show that the estimator of σ_r^2 used here (28) is a uniformly best (minimum variance) unbiased estimator.

We need to consider t when $s < r$. When ρ is large, the term containing $(r-t)(t-1)$ will dominate the change in the variance as t changes as shown by curve no. 1 in Figure 4. Curve no. 2 would result from a smaller value of ρ.

Curve no. 3 shows the case where a local minimum may occur. The only term that contributes toward increasing $V(\hat{\sigma}_r^2)$ as t increases is $(t-1)(r-t)$. This term's greatest increase occurs from $t = 1$ to $t = 2$. If a local minimum occurs, the variance would have to at least decrease from $t = 1$ to $t = 2$ and then increase for some $t > 2$. This case would be fairly unlikely since the incremental increases from $(t-1)(r-t)$ become less as t increases. This term also has the largest denominator in (53).

When ρ is small, the first term of (53) dominates the variance. This term is a monotonically decreasing function of t. This case is shown by curve no. 4.

When $\rho = 0$, the variance of $\hat{\sigma}_r^2$ is strictly a monotonically decreasing function of t as shown by curve no. 5 in Figure 4.

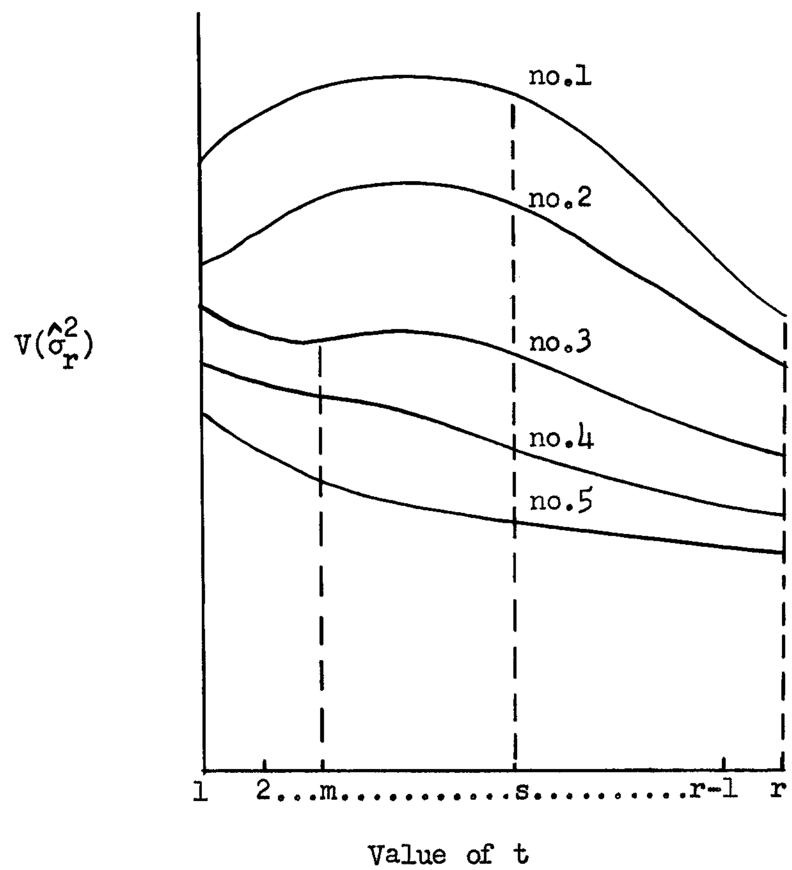

Figure 4. $V(\hat{\sigma}_r^2)$ versus t for various values of ρ

The following procedure will lead to the value of t which minimizes the variance of $\hat{\sigma}_r^2$. Denote by $V'(1)$ the derivative of $V(\hat{\sigma}_r^2)$ at $t = 1$.

(i) If s=r, choose t=r.

(ii) If $s < r$ and $V'(1) \geq 0$, try t=1 and t=s. Select the one which minimizes the variance of $\hat{\sigma}_r^2$.

(iii) If $s < r$ and $V'(1) < 0$, determine if V has a local minimum in the range $1 \leq t < (r+1)/2$. If not, set t=s.

(iv) If $s < r$, $V'(1) < 0$, and if V has a local minimum in the range $1 \leq t \leq (r+1)/2$; denote by m the integer on the t scale closest to this minimum variance.

(a) if $s \leq m$, set t=s.

(b) If $s > m$, try t=m and t=s. Select the one which minimizes the variance of $\hat{\sigma}_r^2$.

Case (iv)(b) is the only case in which t may not be equal to 1 or s. But, selecting t=s instead of t=m would probably have little affect on the variance since both m and s would generally be fairly close to each other when m is better.

A good operational rule for determining t is: try t=1 (u=0) and t=s (u=s) and select the one which minimizes the variance of $\hat{\sigma}_r^2$ as given by (53). It should be pointed out that if t=1 gives the minimum variance, only N-s observations are used and the design is balanced with r rows and c-1 columns. Unless s=r, the use of t=s results in an unbalanced design with all N observations used.

The selection of t is generally not too critical as illustrated by the following problem. The effect due to t is greatest for a small N. The case was investigated where N=25 and r=10 for values of t=1 and 5 where ρ = 1/2, 1,2, and 10. The variances for these cases are summarized in Table 3.

Table 3. $V(\hat{\sigma}_r^2)$ for various values of t and ρ, N=25 and r=10

ρ	t	$V(\hat{\sigma}_r^2)/(2\sigma^4)$
0.5	1	.1389
0.5	5	.1059
1.0	1	.2778
1.0	5	.2381
2.0	1	.7222
2.0	5	.6761
10.0	1	12.28
10.0	5	12.51

To summarize: for the estimator given by (28) with fixed N and r, the design of the type $n_{ij} = 0$ or n which minimizes the variance of $\hat{\sigma}_r^2$ consists of c-1 columns with r rows for each column and one column with u of the r rows, $0 \leq u \leq s \leq r$, with n = 1, i.e. one observation in each of the occupied cells.

Up to this point, r has been kept fixed. The question now remains, having found the form of the optimal design for a given r, of the type $n_{ij} = 0$ or n, what is the best value to choose for r? This final question will now be investigated.

For the form of the optimal design shown in Figure 3, the exact variance of $\hat{\sigma}_r^2$ as given by (52) is easy to find. But, it is still difficult to work with this function to determine the values of r and u which minimize the variance since c and s are functions of r. Thus, an approximate variance will be used which depends on only one design parameter, c_o. A value, $\tilde{r}$, can be found which minimizes the approximate variance of $\hat{\sigma}_r^2$. Then, integers above and below $\tilde{r}$ will be substituted in the formula for the exact variance (52), using the optimal allocation for each, until the value of r is found which minimizes the exact variance of $\hat{\sigma}_r^2$. This scheme will now be developed.

Due to the near balance of the optimal design, the latent roots corresponding to the individual degrees of freedom for R^* with $n = 1$ do not differ by more than $\sigma_r^2/(N-c)$. A good approximation for the variance of $\hat{\sigma}_r^2$ is obtained by using the average latent root $(\sigma_e^2 + \sigma_{rc}^2 + c_o\sigma_r^2)/(N-c)$ where $c_o = (N-c)/(r-1)$. Then $(r-1)R^*$ is distributed approximately as $(\sigma_e^2 + \sigma_{rc}^2 + c_o\sigma_r^2)\chi^2$ with r-1 degrees of freedom. Then, the variance of the estimator of σ_r^2 given by (28) is approximately

$$V(\hat{\sigma}_r^2) \approx \frac{2(\sigma_e^2 + \sigma_{rc}^2 + c_o\sigma_r^2)^2}{c_o^2\,(r-1)} + \frac{2(\sigma_e^2 + \sigma_{rc}^2)^2}{c_o^2\,(N-r-c+1)} \quad .$$

Let

$$\sigma^2 = \sigma_e^2 + \sigma_{rc}^2$$

and

$$\rho = \sigma_r^2/\sigma^2 \quad .$$

Then the approximate variance may be written as

$$V(\hat{\sigma}_r^2) \approx \frac{2\sigma^4}{c_o^2}\left[\frac{(1 + c_o\rho)^2}{r-1} + \frac{1}{N-r-c+1}\right] .$$

From (29)

$$(r-1) = (N-c)/c_o \text{ where } n = 1 \quad .$$

Substituting this result into the above expression gives

$$V(\hat{\sigma}_r^2) \approx \frac{2\sigma^4\,(1 + 2c_o\rho + c_o^2\rho^2 - 2\rho - c_o\rho^2)}{(c_o-1)(N-c)} \quad . \tag{54}$$

From (41)

$$N = r(c-1) + s$$

where $1 \leq s \leq r$. Substituting the limits of s into this give

$$r(c-1) + 1 \leq N \leq rc.$$

Substituting these limits into (29) gives

$$c-1 \leq c_o \leq c$$

or

$$c_o \leq c \leq c_o + 1 \quad . \tag{55}$$

Choosing the mid-point of this interval, $c = c_o + 1/2$, and substituting this central value of c into (54) gives

$$V(\hat{\sigma}_r^2) \approx \frac{2\sigma^4 \,(1 + 2c_o\rho + c_o^2\rho^2 - 2\rho - c_o\rho^2)}{(c_o-1)(N-c_o - 1/2)} \quad . \tag{56}$$

The approximate value $\tilde{c}_o$ of c_o which minimizes (54) is found by setting the derivative of (56) with respect to c_o equal to zero which gives

$$\begin{bmatrix} (N - \tilde{c}_o - 1/2)(\tilde{c}_o - 1)(2\rho + 2\rho^2\tilde{c}_o - \rho^2) \\ - (1 + 2\rho\tilde{c}_o + \rho^2\tilde{c}_o^2 - 2\rho - \rho^2\tilde{c}_o)(N - 2\tilde{c}_o + 1/2) \end{bmatrix} = 0 \quad .$$

This reduces to

$$\tilde{c}_o^2 \left[\rho^2(N - 1/2) + 2\rho \right] + \tilde{c}_o \left[- 2\rho^2(N - 1/2) + 2 - 4\rho \right]$$

$$+ \left[\rho^2(N - 1/2) - (N - 1/2) + 2\rho - 1 \right] = 0 \quad .$$

Solving for $\tilde{c}_o$ gives

$$\tilde{c}_o = \frac{\rho^2(N - 1/2) + 2\rho - 1 + \sqrt{\rho^2(N - 1/2)(N - 3/2) + 2\rho\,(N - 3/2) + 1}}{\rho^2\,(N - 1/2) + 2\rho} \quad . \tag{57}$$

The negative radical is not a solution because this gives values of $c_o < 1$ which are impossible.

Since the result in (57) is only to be used as a first order approximation, replacing (N - 1/2) under the radical by (N - 3/2) still results in a good approximation for c_o. Then, (57) becomes

$$\tilde{c}_o = \frac{\rho(N - 1/2) + (N - 1/2) + 1}{\rho\ (N - 1/2) + 2} \quad . \tag{58}$$

The limiting value of $\tilde{c}_o$ as N increases is $1 + 1/\rho$. For ρ small, this limit may not be too useful.

From (29)

$$r = (N-c + c_o)/c_o \quad .$$

To obtain a first approximation for the value of r which minimizes the variance of $\hat{\sigma}_r^2$, substitute the value of $\tilde{c}_o$ from (58) into the above equation giving

$$\tilde{r} = (N - \tilde{c} + \tilde{c}_o)/\tilde{c}_o \tag{59}$$

where $\tilde{c}$ is the smallest integer greater than or equal to $\tilde{c}_o$. This value of r will serve as a starting point from which to find the value of r which minimizes the exact variance of $\hat{\sigma}_r^2$.

Now, to determine the value of r which minimizes the exact variance of $\hat{\sigma}_r^2$ as given by (52), consider as a first approximation for r, the value r given by (59) which minimizes the approximate variance of $\hat{\sigma}_r^2$. Let r_1 be the largest integer less than r and let r_2 be the smallest integer greater than or equal to r. Obtain the values of u_1 and u_2 which minimize the exact variance of $\hat{\sigma}_r^2$ for r_1 and r_2.

(i) If the exact variance of $\hat{\sigma}_r^2$ as given by (52) for r_1 is less than the variance obtained with r_2, then choose r_3 as the next integer less than r_1. Find the value u_3 which minimizes the exact variance of $\hat{\sigma}_r^2$ for r_3. Continue in this manner until the variance increases. Use the value of r which minimizes the variance of $\hat{\sigma}_r^2$.

(ii) If the exact variance of $\hat{\sigma}_r^2$ as given by (52) for r_2 is less than the variance obtained with r_1, then choose r_3 as the next integer larger than r_2. Find the value u_3 which minimizes the exact variance of $\hat{\sigma}_r^2$ for r_3. Continue in this manner until the variance increases. Use the value of r which minimizes the variance of $\hat{\sigma}_r^2$.

(iii) If the exact variances of $\hat{\sigma}_r^2$ using either r_1 and u_1 or r_2 and u_2 are equal, then choose the one which used the smaller $N^* = N - s + u$.

It has been suggested that one should evaluate these designs on the basis of maximum information per observation used, where the number of observations used is N^*. This is equivalent to minimizing the reciprocal of the information per observation used, i.e. $N^* V(\hat{\sigma}_r^2)$. In this case, it appears that one would usually use a balanced design with c-1 or c columns and r rows, such that $(c-1)r = N-s$ or $cr = N$. A major exception to this procedure would occur if the optimal r is such that $c = 2$ and $2r > N$. In this case, an unbalanced design would be necessary unless a sample size larger than N is used. It is hoped to examine this method of estimating optimal designs in a future investigation.

To obtain the optimal design some previous knowledge of the relative sizes of σ_r^2 and σ^2 are required. It will be shown in Section 6.6 that the value actually used for ρ can differ considerably from the true value without appreciably increasing the variance of $\hat{\sigma}_r^2$.

Comparisons of the approximate variance given by (56) and the exact variance given by (52) using $t = u = r/2$ are shown in Table 4 for several cases. The percent differences were computed in the following manner

$$\text{Percent difference} = \frac{\text{(Exact variance) - (Approximate Variance)}}{\text{Exact variance}} \times 100\%.$$

The percent differences in Table 4 were computed using six decimal places for the variances although the variances are shown with four decimals.

The first three groups in Table 4 with one full and one half-full column of observations show for each ρ that the percent differences do not increase appreciably as N decreases. The fourth group shows that adding one full column of observations and keeping N=30 the same, the approximation becomes closer for each ρ. The last two groups show that changing the half column to 2/3 or 1/4 columns of observations and keeping N=30 the same, the approximation is closer for each ρ. When N/r is an integer, a balanced design is obtained and the approximate variance is then identically equal to the exact variance. Table 4 indicates that a design consisting of one full and one half-full column of observations apparently leads to the worst approximations.

Table 4. Comparison of approximate and exact variances of $\hat{\sigma}_r^2$

N	r	N/r	ρ	Exact variance ($x\sigma^4$)	Approximate variance ($x\sigma^4$)	Percent difference
75	50	1.5	0	.0559	.0559	0
75	50	1.5	1	.1561	.1515	2.94
75	50	1.5	2	.3472	.3288	5.30
75	50	1.5	4	1.0017	.9282	7.34
75	50	1.5	8	3.4006	3.1073	8.62
45	30	1.5	0	.0963	.0963	0
45	30	1.5	1	.2662	.2583	2.94
45	30	1.5	2	.5896	.5583	5.31
45	30	1.5	4	1.6972	1.5719	7.38
45	30	1.5	8	5.7556	5.2559	8.68
30	20	1.5	0	.1508	.1508	0
30	20	1.5	1	.4110	.3989	2.94
30	20	1.5	2	.9059	.8576	5.33
30	20	1.5	4	2.5998	2.4065	7.44
30	20	1.5	8	8.8039	8.0336	8.75
30	12	2.5	0	.0509	.0509	0
30	12	2.5	1	.3884	.3808	1.96
30	12	2.5	2	1.1044	1.0744	2.72
30	12	2.5	4	3.6723	3.5525	3.26
30	12	2.5	8	13.3513	12.8722	3.59
30	18	1.67	0	.1104	.1104	0
30	18	1.67	1	.3808	.3709	2.60
30	18	1.67	2	.9064	.8667	4.38
30	18	1.67	4	2.7228	2.5642	5.82
30	18	1.67	8	9.4172	8.7827	6.74
30	24	1.25	0	.3286	.3286	0
30	24	1.25	1	.5684	.5583	1.78
30	24	1.25	2	1.0020	.9618	4.01
30	24	1.25	4	2.4510	2.2902	6.56
30	24	1.25	8	7.6754	7.0326	8.37

From (58), optimal designs with one full and one half-full column of observations arise when ρ is approximately two. For, N=30, N/r = 1.5, and ρ = 2 the difference between the approximate and exact variance is 5.33%. It appears that the approximation for $V(\hat{\sigma}_r^2)$ will generally be close to the exact variance.

To summarize the results of this section:

(a) The estimator, $\hat{\sigma}_r^2$, considered in (28) was obtained by equating the mean squares to their expected values giving a unique unbiased estimator where the mean squares were obtained by the method of fitting constants with the row mean squares adjusted for columns and interaction adjusted for both rows and columns. As indicated earlier, other estimators may exist which used in connection with their optimal designs could give smaller variances under certain conditions. Even for the estimator selected, it was not possible to make an exhaustive study of this problem. In order to limit the size of this investigation to a reasonable size, it was decided to consider only connected designs of the type $n_{ij} = 0$ or n.

(b) It was proved that the design of this type which minimizes the variance of $\hat{\sigma}_r^2$ for a fixed N and r is obtained with $n_{ij} = 0$ or 1. The results of this study do not preclude the possibility that a more general design would be better.

(c) N can be represented as follows

$$N = r(c-1) + s \quad , \qquad 0 < s \leq r \quad .$$

The optimal design consists of one observation in each cell for c-1 columns and r rows and one observation in $u(0 \leq u \leq s \leq r, u \neq 1)$ cells of one column. Hence, the number of observations actually used would be $N^* = r(c-1) + u$. A procedure was developed to determine u. In most cases u will be either 0 or s.

(d) The exact variance for $\hat{\sigma}_r^2$ is relatively easy to find for the design given by (c) due to the near balance.

(e) To find the value of r which minimizes the exact variance of $\hat{\sigma}_r^2$, it is easier to obtain a first approximation, $\tilde{r}$, which minimizes the approximate variance. Then, values of r below and values of r above $\tilde{r}$ are used to compute the exact variance. For each value of r tried, the value of u which minimizes the variance of $\hat{\sigma}_r^2$ is determined. The combination of r and u which minimizes the exact variance of $\hat{\sigma}_r^2$ is found in this manner.

(f) All of the results in this section apply to the estimation of σ_c^2 by simply interchanging r and c and by interchanging σ_r^2 and σ_c^2.

6.6 Effect of Improper Choice of ρ in Determining Optimal Design for $\hat{\sigma}_r^2$

In order to determine the value of r which minimizes the variance of $\hat{\sigma}_r^2$ according to the procedure given in Section 6.5, some knowledge of ρ is required. Denote by ρ', the incorrect value of ρ used to determine the optimal design. Let V denote the approximate variance of $\hat{\sigma}_r^2$ given by (56) for the optimal design based on the true value of ρ. Let V' denote the approximate variance of $\hat{\sigma}_r^2$ for the optimal design based on ρ'. Table 5 shows how the variance of $\hat{\sigma}_r^2$ changes when the value used for ρ to determine the optimal design differs from the true value of ρ by a factor of two. R.E. is the relative efficiency of the design based on ρ' to the optimal design based on ρ where

$$R.E. = (V/V') \times 100\%.$$

Table 5. Effect on the approximate variance of $\hat{\sigma}_r^2$ due to using improper choices of ρ in determining optimal designs

N	ρ	$\tilde{c}_o$	V/σ^4	ρ'	$\tilde{c}_o'$	V'/σ^4	R.E. (%)
30	.25	4.04	.0849	.125	6.01	.0915	92.8
30	.25	4.04	.0849	.500	2.70	.0938	90.5
30	1.0	1.90	.3631	.50	2.70	.3946	92.0
30	1.0	1.90	.3631	2.00	1.47	.3994	90.9
30	4.0	1.24	2.265	2.00	1.47	2.402	94.3
30	4.0	1.24	2.265	8.00	1.12	2.411	93.9
100	.25	4.67	.0225	.125	7.82	.0248	90.7
100	.25	4.67	.0225	.500	2.90	.0250	90.0
100	1.0	1.97	.1026	.50	2.90	.1125	91.2
100	1.0	1.97	.1026	2.00	1.49	.1129	90.9
100	4.0	1.25	.6514	2.00	1.49	.6914	94.2
100	4.0	1.25	.6514	8.00	1.12	.6963	93.6

For example, consider the first case in Table 5. For N=30 and ρ=.25 the optimal design according to the procedure developed in Section 6.5 would have approximately 4 columns ($\tilde{c}_o$) and approximately N/4 rows. However, if ρ is unknown and a guess of ρ=.125 is used, the procedure in Section 6.5 would give a design with approximately 6 columns ($\tilde{c}_o$) and approximately N/6 rows. The relative efficiency of this design to the optimal design is 92.8%, where the variances are computed from (56).

Table 5 indicates that the standard deviation of $\hat{\sigma}_r^2$ from the optimal design based on ρ would vary from about 95 to 98% of the standard deviation of $\hat{\sigma}_r^2$ obtained from the optimal design based on 2ρ or $\rho/2$. Even though the procedure for finding an optimal design requires some knowledge of ρ, the value used for ρ can vary considerably from the true value without seriously affecting the variance of the estimator of σ_r^2.

Due to the closeness of the approximate variance to the exact variance, the relative efficiencies based on exact variances will be close to those shown in Table 5 which are based on approximate variances. For example, suppose ρ=1.0 and N=30. The optimal design has the following parameters: r=15, c=2 and u=15. The exact variance of $\hat{\sigma}_r^2$ is $0.3571\sigma^4$. Suppose a guess of ρ=2.0 is used. The optimal design based on ρ=2.0 consists of r=19, c=2 and u=11. The exact variance of $\hat{\sigma}_r^2$ is $0.3939\sigma^4$. The relative efficiency of this design to the optimal design based on exact variances is 90.7% as compared to 90.9% shown in Table 5 using approximations.

6.7 Estimation of σ_r^2 or σ_c^2 when $\sigma_{rc}^2 = 0$

When it is assumed that there are no interaction effects ($\sigma_{rc}^2 = 0$), the model given by (1) for the two-way crossed classification is

$$y_{ijk} = \mu + r_i + c_j + e_{ijk} \quad . \tag{60}$$

Now, no restriction is placed on the number of observations per cell other than $n_{ij} \geq 0$ and $\sum n_{ij} = N$.

The analysis of variance is given in Table 6 for a connected design, using the method of fitting constants.

Table 6. Analysis of variance for two-way crossed classifications with unequal numbers per cell and no interaction

Source of variation	DF	MS	E(MS)
Columns	c-1	C	$\sigma_e^2 + c_1\sigma_r^2 + r_o\sigma_c^2$
Rows (adjusted for columns)	r-1	R^*	$\sigma_e^2 + c_o^*\sigma_r^2$
Rows	r-1	R	$\sigma_e^2 + r_1\sigma_c^2 + c_o\sigma_r^2$
Columns (adjusted for rows)	c-1	C^*	$\sigma_e^2 + r_o^*\sigma_c^2$
Error	N-r-c+1	E	σ_e^2
Total	N-1		

If only one column were used and all observations in the i^{th} row of the c-column design were put in the i^{th} row of the 1-column design, $E(MSR_1)$ would be the same as E(MSR) above except that $r_1 = 0$. That is, $E(MSR_1) = \sigma_e^2 + c_o\sigma_r^2$, where $c_o = (N-\sum n_{i.}^2/N)/(r-1)$. The best possible situation for estimating σ_r^2 with the c-column design would appear to occur when $\sigma_c^2 = 0$. That is, only σ_r^2 and σ_e^2 need be estimated and only σ_r^2 and σ_e^2 appear in the variance of $\hat{\sigma}_r^2$. But in this most favorable case, MSR will have identically the same distribution as MSR_1. Also, the variance of $\hat{\sigma}_e^2$ will be greater with c columns unless SSC^* and SSE are pooled, in which case the variance of $\hat{\sigma}_e^2$ will be the same in the two cases. Since under the most favorable situation of $\sigma_c^2 = 0$, the c-column design is no better than the 1-column design, it would appear that under the general condition of $\sigma_c^2 \geq 0$, the 1-column design would be better.

A more formal approach would be to prove that under the assumption that $\sigma_c^2 = 0$

$$V(MSR)/c_o^2 \leq V(MSR^*)/c_o^{*2} ,$$

where

$$c_o^* = \left[N - \sum_j \sum_i \frac{n_{ij}^2}{n_{\cdot j}} \right] / (r-1) .$$

Since

$$\sum_j \frac{n_{ij}^2}{n_{\cdot j}} \geq \frac{n_{i\cdot}^2}{N} ,$$

$c_o \geq c_o^*$. Hence, it is easy to show that the coefficients of σ_e^4 and $\sigma_e^2\sigma_r^2$ are never greater for V(MSR) than for V(MSR*). The problem is more complicated for the coefficients of σ_r^4. In this case it is required to prove that

$$\frac{\sum_i n_{i\cdot}^2 - \frac{2}{N} \sum_i n_{i\cdot}^3 + \frac{1}{N^2} \left[\sum_i n_{i\cdot}^2 \right]^2}{c_o^2}$$

is less than or equal to

$$\frac{\sum_i n_{i\cdot}^2 - 2 \sum_i \sum_j \frac{n_{i\cdot} n_{ij}^2}{n_{\cdot j}} + \sum_i \sum_\ell \left[\sum_j \frac{n_{ij} n_{\ell j}}{n_{\cdot j}} \right]^2}{c_o^{*2}} .$$

The proof of this inequality has not been attempted here, since the earlier argument that the 1-column design is better than or equivalent to the c-column design seemed sufficient.

Assuming we have established that, if $\sigma_{rc}^2 = 0$, the 1-column design is best, we next proceed to find the best allocation in this column. This problem was solved by Crump [1954] where the r rows can be identified with his a classes. The optimal design for this situation was presented in Section 4.0.

The results of this section apply to the estimation of σ_c^2, with $\sigma_{rc}^2 = 0$, by interchanging r and c.

6.8 Estimation of $\sigma_r^2/(\sigma_e^2 + \sigma_{rc}^2)$ or $\sigma_c^2/(\sigma_e^2 + \sigma_{rc}^2)$

The problem of estimating $\sigma_r^2/(\sigma_e^2 + \sigma_{rc}^2)$ will now be studied. Consider the analysis of variance given in Table 2 which is obtained by the method of fitting constants for the model given by (1). Alternative analyses of variance would lead to alternative estimators of $\rho = \sigma_r^2/(\sigma_e^2 + \sigma_{rc}^2)$. The investigation here will be limited to obtaining an estimator of ρ from this analysis of variance.

For a design with $n_{ij} = 0$ or n, an estimator of $\sigma_r^2/(\sigma_e^2 + n\sigma_{rc}^2)$ is readily obtained by considering the ratio $F' = R^*/I^*$ in Table 2. Setting $n_{ij} = 0$ or 1 leads to a simple estimator of $\rho = \sigma_r^2/(\sigma_e^2 + \sigma_{rc}^2)$ based on the ratio $F' = R^*/I^*$. For the more general design with $n_{ij} = 0$ or n the estimator of ρ is far more complicated.

R^* can be divided into its r-1 orthogonal sum of squares giving

$$(r-1)R^* = \sum_{i=1}^{r-1} (\sigma^2 + c_i\sigma_r^2)\chi_i^2 \quad , \tag{61}$$

where $\sigma^2 = \sigma_e^2 + \sigma_{rc}^2$ and each χ_i^2 has one degree of freedom. Also,

$$(N-r-c+1)\ I = \sigma^2\chi^2$$

where χ^2 has (N-r-c+1) degrees of freedom. Then,

$$F' = R^*/I = \frac{(N-r-c+1)}{(r-1)} \sum_{i=1}^{r-1} \frac{(\sigma^2 + c_i\sigma_r^2)\chi_i^2}{\sigma^2\chi^2}$$

or

$$F' = \frac{1}{r-1} \sum_{i=1}^{r-1} (1 + c_i\rho)\, F_i \tag{62}$$

where F_i is a F variate with 1 and N-r-c+1 degrees of freedom. The expected value of F with m_1 and m_2 degrees of freedom is

$$E(F) = m_2/(m_2 - 2) \quad , \qquad m_2 > 2 \quad .$$

Therefore

$$E(F_i) = (N-r-c+1)/(N-r-c-1) \; .$$

Substituting this result into (62) gives

$$E(F') = \frac{(N-r-c+1)}{(N-r-c-1)} \left[1 + \frac{\rho}{(r-1)} \sum_{i=1}^{r-1} c_i \right] . \tag{63}$$

From (61) the expected value of R^* is

$$E(R^*) = \sigma^2 + \frac{\sigma_r^2}{r-1} \sum_{i=1}^{r-1} c_i \tag{64}$$

and from Table 2

$$E(R^*) = \sigma^2 + c_o\sigma_r^2 \quad . \tag{65}$$

Equating (64) and (65) gives

$$c_o = \frac{1}{r-1} \sum_{i=1}^{r-1} c_i \quad . \tag{66}$$

Substituting this result into (63) gives

$$E(F') = \frac{(N-r-c+1)}{(N-r-c-1)} \; (1 + c_o\rho) \; .$$

This result leads to the following unbiased estimator of ρ based on F'

$$\hat{\rho} = \frac{1}{c_o}\left[\frac{(N-r-c-1)}{(N-r-c+1)} \; F' - 1\right] \tag{67}$$

where c_o is given by (29)

$$c_o = (N-c)/(r-1) \; . \tag{68}$$

The variance of $\hat{\rho}$ is

$$V(\hat{\rho}) = \frac{(N-r-c-1)^2}{(N-r-c+1)^2} \frac{V(F')}{c_o^2} \; . \tag{69}$$

From (62), the variance of F' is

$$V(F') = \frac{V(F_i)}{(r-1)^2}\sum_{i=1}^{r-1} (1 + c_i\rho)^2$$

$$+ \frac{\mathrm{Cov}(F_i,F_{j\neq i})}{(r-1)^2} \sum_{i=1}^{r-1} \sum_{j\neq i}^{r-1} (1 + c_i\rho)(1 + c_j\rho) \; . \tag{70}$$

The moments of F with m_1 and m_2 degrees of freedom are

$$E(F^k) = \frac{(m_1/2 + k - 1)!}{(m_1/2 - 1)!} \frac{(m_2/2 - k - 1)!}{(m_2/2 - 1)!} \left[\frac{m_2}{m_1}\right]^k$$

where $k < m_2/2$. Then, the variance of F is

$$V(F) = E(F^2) - E(F)^2$$

or

$$V(F) = \frac{2\, m_2^2\, (m_1 + m_2 - 2)}{(m_2 - 2)^2\, (m_2 - 4)m_1} \; , \qquad m_2 > 4 \; . \tag{71}$$

In our case, F_i has $m_1 = 1$ and $m_2 = N-r-c+1$ degrees of freedom.

Thus,

$$V(F_i) = \frac{2\, m_2^2\, (m_2 - 1)}{(m_2 - 2)^2\, (m_2 - 4)} \; . \tag{72}$$

Crump [1954, p. 39] gives the covariance between two F variates which have orthogonal χ^2 with single degrees of freedom in the numerator and a common denominator of χ^2/m_2,

$$\mathrm{Cov}(F_i, F_j) = \frac{2\, m_2^2}{(m_2 - 2)^2 (m_2 - 4)} \quad . \tag{73}$$

Substituting from (72) and (73) into (70) gives

$$V(F') = \frac{2\, m_2^2\, (r-1)^{-2}}{(m_2 - 2)^2 (m_2 - 4)} \left[\begin{array}{l} (m_2-1) \sum\limits_{i=1}^{r-1} (1 + c_i\rho)^2 \\ + \sum\limits_{i=1}^{r-1} \sum\limits_{j \neq 1}^{r-1} (1 + c_i\rho)(1 + c_j\rho) \end{array} \right] .$$

Substituting this result into (69) gives

$$V(\hat{\rho}) = \frac{2\,(r-1)^{-2}}{c_o^2 (m_2 - 4)} \left[\begin{array}{l} (m_2-2) \sum (1 + c_i\rho)^2 \\ + \left[\sum (1 + c_i\rho) \right]^2 \end{array} \right] \tag{74}$$

for $m_2 > 4$. It follows from (66) that $\sum (1 + c_i\rho)$ is a constant for any particular c

$$\sum_{i=1}^{r-1} (1 + c_i\rho) = (r-1)(1 + c_o\rho) \quad . \tag{75}$$

Thus,

$$\sum_{i=1}^{r-1} (1 + c_i\rho)^2 = \sum (1 + c_i\rho)^2 - \frac{\left[\sum (1 + c_i\rho) \right]^2}{r-1} + \frac{\left[\sum (1 + c_i\rho) \right]^2}{r-1}$$

or

$$\sum_{i=1}^{r}(1 + c_i\rho)^2 = \rho^2\sum(c_i-c_o)^2 + (r-1)(1 + c_o\rho)^2 \quad . \tag{76}$$

Substituting from (75) and (76) into (74) gives

$$V(\hat{\rho}) = \frac{2(r-1)^{-2}}{c_o^2\,(m_2-4)}\left[\begin{array}{l}(m_2-2)\,\rho^2\sum(c_i - c_o)^2 \\ + (m_2-2)(r-1)(1 + c_o\rho)^2 \\ + (r-1)^2\,(1 + c_o\rho)^2\end{array}\right]. \tag{77}$$

$V(\hat{\rho})$ is minimized for fixed N and r by making $\sum(c_i - c_o)^2$ small and m_2 and c_o large. Since

$$m_2 = N-r-c+1$$

and

$$c_o = (N-c)/(r-1) \quad ,$$

these are both maximized when c is minimized. Thus, the situation here is analogous to that encountered in Section 6.5 for the estimation of σ_r^2. In some cases, it will be desirable to discard observations. N can be represented as

$$N = r(c-1) + s \quad , \qquad 0 < s \leq r \quad .$$

As seen in Section 6.5, for the form of a variance function as shown in (77), the variance of $\hat{\rho}$ is minimized by using c-1 columns with r rows in common and one column with u ($0 \leq u \leq s \leq r$, $u \neq 1$) of the r rows. The value of u=1 does not occur because in this case s observations are completely confounded with columns and can be discarded without changing the variance of $\hat{\rho}$. The form of the optimal design for minimizing the variance of $\hat{\rho}$ is shown in Figure 3. As before, the number of observations used is $N^* = N-s+u$.

Up to this point, r has been kept fixed. The question now remains, having found the form of the design with $n_{ij} = 0$ or 1 which minimizes the variance of $\hat{\rho}$, what is the best value to choose for r? This final question will now be investigated. Due to the near balance of the optimal design, the variance of $\hat{\rho}$ is fairly easy to obtain. But it is still difficult to work with this variance in order to determine the value of r and u which minimizes the variance of $\hat{\rho}$. Thus, as was done for estimating σ_r^2, an approximate variance of $\hat{\rho}$ will be used to determine an approximate value, $\tilde{r}$, which minimizes the exact variance of $\hat{\rho}$. Then, integers above and below $\tilde{r}$ are tried until the value of r and u is determined which minimizes the exact variance of $\hat{\rho}$.

$F'/(1 + c_o\rho)$ is distributed as F when the same columns are sampled for each row. $F'/(1 + c_o\rho)$ is distributed approximately as F when approximately the same columns are sampled for each row as is the case here since it was shown that the variance of $\hat{\rho}$ is minimized when the number of columns sampled for some rows is c-1 and c for the remaining rows. Under these conditions the sum of squares for rows is approximately distributed as $(\sigma^2 + c_o\sigma_r^2)\chi^2$. This is the same approximation used in Section 6.5 for obtaining an approximate variance of $\hat{\sigma}_r^2$. Thus,

$$F' \approx (1 + c_o\rho) F \quad .$$

When N is large, the variance of F from (71) is approximately

$$V(F) \approx 2 (m_1 + m_2)/m_1m_2 \quad .$$

Hence,

$$V(F') \approx 2(1 + c_o\rho)^2(m_1 + m_2)/m_1m_2 \ .$$

Since $m_1 = r-1$ and $m_2 = N-c-r+1$,

$$V(\hat{\rho}) \approx \frac{2\ (1 + c_o\rho)^2\ (N - c)}{c_o^2(r-1)\left[(N-c) - (r-1)\right]} \tag{78}$$

From (29)

$$(r-1) = (N-c)/c_o \ .$$

Substituting this result in (78) gives

$$V(\hat{\rho}) \approx 2\ (1 + c_o\rho)^2/(c_o-1)(N-c) \ . \tag{79}$$

From (55), $c_o \leq c \leq c_o + 1$. Substituting the mid-point of this interval, $c = c_o + 1/2$, into (79) gives

$$V(\hat{\rho}) \approx 2\ (1 + c_o\rho)^2/(c_o - 1)(N - c_o - 1/2) \ . \tag{80}$$

The value, $\tilde{c}_o$, of c_o which minimizes (80) is obtained by setting the derivative of (80) with respect to c_o equal to zero, which gives

$$\tilde{c}_o = \frac{2\rho\ (N - 1/2) + (N - 1/2) + 1}{\rho\ (N - 1/2) + \rho + 2} \ . \tag{81}$$

From (29)

$$r = (N - c + c_o)/c_o \ . \tag{82}$$

To obtain a first approximation for r, substitute the value of $\tilde{c}_o$ from (81) into (82) giving

$$\tilde{r} = (N - \tilde{c} + \tilde{c}_o)/\tilde{c}_o \tag{83}$$

where $\tilde{c}$ is the smallest integer greater than or equal to $\tilde{c}_o$. This value, $\tilde{r}$, will serve as a starting point from which to find the value of r which minimizes the exact variance of $\hat{\rho}$. To obtain the value of r requires some previous knowledge of ρ since it appears as a nuisance parameter in the variance of $\hat{\rho}$. It will be shown in Section 6.9 that the value actually used for ρ can differ considerably from the true value of ρ without appreciably increasing the variance of $\hat{\rho}$.

The exact variance of $\hat{\rho}$ will now be obtained for the form of the optimal design obtained previously in this section with $n_{ij} = 0$ or 1. This design consists of c-1 columns with r rows and one column with u rows. The occupied cells contain one observation. The form of the design is illustrated in Figure 3 and the analysis of variance showing the orthogonal parts of the row sum of squares is given in Table 7. From Table 7

$$R^* = \frac{(r-t-1)R_1^* + (t-1)R_2^* + R_3^*}{r-1} , \qquad \begin{array}{l} t=u \text{ for } u \geq 2 \\ t=1 \text{ for } u = 0 . \end{array}$$

Since R_1^* and R_3^* have the same expected values, they can be pooled giving

$$R_4^* = \left[(r-t-1)R_1^* + R_3^*\right] / (r-t) .$$

Substituting these results into (67) gives

$$\hat{\rho} = \frac{(N-r-c-1)}{(N-c)(N-r-c+1)} \left[(t-1)\frac{R_2^*}{I} + (r-t)\frac{R_4^*}{I} \right] - \frac{r-1}{N-c} . \tag{84}$$

Let $R_2^*/I = F_2'$ and $R_4^*/I = F_4'$. Then $\hat{\rho}$ becomes

$$\hat{\rho} = \frac{(N-r-c-1)}{(N-c)(N-r-c+1)} \left[(t-1)F_2' + (r-t)F_4' \right] - \frac{r-1}{N-c} . \tag{85}$$

$F_2'/(1 + c\rho)$ is distributed as F with (t-1) and (N-r-c+1) degrees of freedom and $F_4'/[1 + (c-1)\rho]$ is distributed as F with (r-t) and (N-r-c+1) degrees of freedom. The variance of the unbiased estimator, $\hat{\rho}$, based on the ratio $F' = R^*/I$ is

$$V(\hat{\rho}) = \left(\frac{N-r-c-1}{(N-c)(N-r-c+1)}\right)^2 \left[\begin{array}{l}(t-1)^2(1+c\rho)^2V(F_2)+(r-t)^2(1+c\rho-\rho)^2V(F_4) \\ + 2(t-1)(r-t)(1+c\rho)(1+c\rho-\rho)Cov(F_2,F_4)\end{array}\right] \quad (86)$$

Table 7. Analysis of variance for optimal design shown in Figure 3

Source of variation	DF	MS	E(MS)
Columns	c-1	C	$\sigma^2 + c_1\sigma_r^2 + r_o\sigma_c^2$
Rows (adjusted for columns)	r-1	R^*	$\sigma^2 + c_o\sigma_r^2$
First r-t rows	r-t-1	R_1^*	$\sigma^2 + (c-1)\sigma_r^2$
Last t rows	t-1	R_2^*	$\sigma^2 + c\sigma_r^2$
First r-t vs last t rows	1	R_3^*	$\sigma^2 + (c-1)\sigma_r^2$
Error	N-r-c+1	I	σ^2
Total	N-1		

where F_2 is distributed as F with (t-1) and (N-r-c+1) degrees of freedom and F_4 is distributed as F with (r-t) and (N-r-c+1) degrees of freedom. Formula (71) gives the variance of F. The numerator mean squares of F_2 and F_4 are orthogonal but they have the same denominator mean square, I. For this case, Crump [1954] gives the covariance of these two F's.

Now, to determine the value of r that minimizes the exact variance of $\hat{\rho}$ as given by (86), consider as a first approximation for r, the value $\tilde{r}$ given by (83). Follow the same procedure of considering integers around $\tilde{r}$ and finding the best value of u for each as given in Section 6.5 leads to the value of r which minimizes the variance of $\hat{\rho}$. The value chosen for c is the largest integer greater than or equal to N/r.

To summarize the results of this section:

(a) Alternative estimators of ρ used in conjunction with their optimal designs may yield smaller variances under certain conditions.

(b) The connected design of the form $n_{ij} = 0$ or 1 which minimizes the exact variance was determined for fixed N and r.
Designs consisting of $n_{ij} = 0$ or 1 lead to a simple estimator for σ_r^2/σ^2 based on the ratio of mean squares. The optimal design consisted of c-1 columns with r rows each and one column with u of the r rows where $N = r(c-1) + s$ and $0 \leq u \leq s \leq r$, $u \neq 1$. The total number of observations used is $N^* = N-s+u$.

(c) The exact variance for $\hat{\rho}$ is relatively easy to find for the form of the design designated by (c) due to the near balance.

(d) To find the value of r that minimizes the exact variance of $\hat{\rho}$, consider as a first approximation to r the value $\tilde{r}$ from (83) which minimizes the approximate variance. Consider integers around $\tilde{r}$ in the manner described in Section 6.5. This leads to the design which minimizes the exact variance of $\hat{\rho}$. For each value of r considered, the best value of u must also be determined.

(e) All of the results of this section apply to the estimation of $\sigma_c^2/\sigma^2 = \sigma_c^2/(\sigma_e^2 + \sigma_{rc}^2)$ by simply interchanging r and c.

6.9 Effect of Improper Choice of ρ in Determining Optimal Design for $\hat{\rho}$

In order to determine the value of r which minimizes the variance of $\hat{\rho}$ according to the procedure given in Section 6.8, some knowledge of ρ is required since ρ is a nuisance parameter in the formula for the variance of $\hat{\rho}$ as given by (86). Denote by ρ' the incorrect value of ρ used to determine the optimal design. In order to simplify the investigation of the effects of using ρ' instead of the true value of ρ in obtaining the optimal design, the approximate variance of $\hat{\rho}$ for large N will be used. The approximate variance of $\hat{\rho}$ for the nearly optimal design based on ρ is obtained from (79)

$$\tilde{V} = 2\ (1 + \tilde{c}_o\rho)^2/(\tilde{c}_o-1)(N-\tilde{c}) \qquad (87)$$

where $\tilde{c}$ is the smallest integer greater than or equal to $\tilde{c}_o$ and $\tilde{c}_o$ is obtained from (81)

$$\tilde{c}_o = \frac{2\rho(N - 1/2) + (N - 1/2) + 1}{\rho(N - 1/2) + \rho + 2} \quad . \qquad (88)$$

Similarly, the approximate variance of $\hat{\rho}$ for the design based on p' is

$$\tilde{V}' = 2\ (1 + \tilde{c}_o'\rho)^2/(\tilde{c}_o'-1)(N-\tilde{c}') \qquad (89)$$

where $\tilde{c}'$ is the smallest integer greater than or equal to $\tilde{c}_o'$ and

$$\tilde{c}_o' = \frac{2\rho'(N - 1/2) + (N - 1/2) + 1}{\rho'(N - 1/2) + \rho' + 2} \quad . \qquad (90)$$

Table 8 shows how the approximate variance of $\hat{\rho}$ changes when the value used for ρ to determine the optimal design differs from the true value by a factor of four. That is, $\rho' = 4\rho$ and $\rho' = \rho/4$. $\widetilde{R.E.}$ is the approximate relative efficiency of the design based on ρ' to the design based on ρ where

$$\widetilde{R.E.} = (V/V') \times 100\%.$$

Table 8. Effect on the approximate variance of $\hat{\rho}$ due to using an improper choice of ρ while determining the optimal design

N	ρ	ρ'	V	V'	R.E.
30	.25	1.00	.1023	.1205	84.9
30	1.00	.25	.5952	.7025	84.7
30	.50	2.00	0.233	0.259	90.0
30	2.00	.50	1.782	1.965	90.7
30	2.00	8.00	1.78	1.84	97.0
30	8.00	2.00	21.37	21.53	99.3
100	.25	1.00	.0267	.0321	82.9
100	1.00	.25	.1650	.2004	82.3
100	.50	2.00	.0626	.0702	89.1
100	2.00	.50	.4949	.5521	89.6
100	2.00	8.00	0.495	0.507	97.7
100	8.00	2.00	5.939	6.040	98.3

For example, consider the first case in Table 8. For N=30 and ρ=.25 the optimal design is obtained with $c_o \approx 4.70$. However, if ρ is unknown and a guess ρ'=1.0 is used to determine the optimal design then $c_o' \approx 2.75$. The approximate relative efficiency of this design to the optimal design is 84.9%.

Table 8 indicates the approximate standard deviation of $\hat{\rho}$ from the optimal design based on ρ would vary from about 90 to almost 100% of the approximate standard deviation of $\hat{\rho}$ obtained from the optimal design based on 4ρ or $\rho/4$. Even though the procedure for finding the optimal design for ρ requires some knowledge of ρ, the value actually used may vary considerably from the true value without seriously affecting the variance of the estimator of $\rho = \sigma_r^2/\sigma^2$.

6.10 Estimation of σ_r^2/σ_e^2 or σ_c^2/σ_e^2 when $\sigma_{rc}^2 = 0$

Consider the special case of estimating σ_r^2/σ_e^2 or σ_c^2/σ_e^2 when $\sigma_{rc}^2 = 0$. Following the same argument used in Section 6.7, it appears that a 1-column design is always better or equivalent to a c-column design. In this case, the best allocation to the rows in this column is solved by Crump [1954], where the r rows can be identified with his a classes. The optimal design for this situation was presented in Section 4.0.

7.0 SIMULTANEOUS ESTIMATION OF σ_r^2 AND σ_r^2/σ^2 OR σ_c^2 AND σ_c^2/σ^2

From (58) the approximate value of c_o which minimizes the variance of $\hat{\sigma}_r^2$ is

$$\tilde{c}_o(\sigma_r^2) = \frac{\rho(N - 1/2) + (N - 1/2) + 1}{\rho(N - 1/2) + 2} \tag{91}$$

where $\rho = \sigma_r^2/\sigma^2$ and $\sigma^2 = \sigma_e^2 + \sigma_{rc}^2$. From (81) the approximate value of c_o which minimizes the variance of $\hat{\rho}$ is

$$\tilde{c}_o(\rho) = \frac{2\rho(N - 1/2) + (N - 1/2) + 1}{\rho(N - 1/2) + \rho + 2} \quad . \tag{92}$$

When N is large, (91) becomes approximately

$$\tilde{c}_o(\sigma_r^2) \approx 1 + 1/\rho \tag{93}$$

and (92) becomes approximately

$$\tilde{c}_o(\rho) \approx 2 + 1/\rho \; . \tag{94}$$

For large N, the optimal design for estimating ρ consists of approximately one more column than the optimal design for estimating σ_r^2.

Table 9 shows the approximate values of c_o for large N for the optimal designs for estimating σ_r^2 and ρ.

Table 9. Approximate values of c_o for the optimal designs for estimating σ_r^2 and ρ (large N)

ρ	Approximate values of c_o	
	Estimating σ_r^2	Estimating ρ
1/4	5	6
1/2	3	4
1	2	3
2	3/2	5/2
4	5/4	9/4

It was seen in Sections 6.6 and 6.9 that the design used could vary considerably from the optimal design without large losses in information. From Table 9, the designs for estimating σ_r^2 and ρ do not differ markedly when $\rho \leq 1$. These cases are well within the limits of discrepancies studied in Sections 6.6 and 6.9.

Allocation for the estimation of σ_r^2 will also give good estimates of ρ and vice versa when $\rho \leq 1$. If estimates of the two values, σ_r^2 and ρ, were equally important, one might consider a design with a c_o value about half-way between $c_o(\sigma_r^2)$ and $c_o(\rho)$ when $\rho \leq 1$. In this way, only a small amount of information would be sacrificed on each estimate.

The remarks of this section apply to the simultaneous estimation of σ_c^2 and σ_c^2/σ^2 by interchanging r and c in the discussion.

8.0 SIMULTANEOUS ESTIMATION OF σ_r^2 AND σ_c^2

The simultaneous estimation of σ_r^2 and σ_c^2 is complicated since one must determine the type of design to use in addition to the allocation for a given design. In general, it has been shown that efficient estimation of σ_r^2 requires a large number of rows and few columns. Similarly, efficient estimation of σ_c^2 requires a large number of columns and few rows.

It was not the original purpose of this thesis to investigate this problem, but a few tentative observations will be presented.

We will consider two possible types of designs, which attempt to sample both a large number of rows and columns.

8.1 L-shaped Design

First, suppose we use N_1 observations primarily to estimate σ_r^2 and N_2 observations primarily to estimate σ_c^2. From Section 6.5, the N_1 observations would consist of a large number of rows and a few columns. The N_2 observations would consist of a large number of columns and few rows. A combination of these two configurations is designated as an L-shaped design and is shown in Figure 5. The total numbers of rows and columns are $r = r_1 + r_2$ and $c = c_1 + c_2$. The N_1 observations contain one column with r_3 $(0 \leq r_3 < r_1)$ empty cells, $N_1 = rc_2 - r_3$. The N_2 observations contain one row with c_3 $(0 \leq c_3 < c_1)$ empty cells, $N_2 = cr_2 - c_3$. The total number of observations is $N = N_1 + N_2 - r_2c_2$.

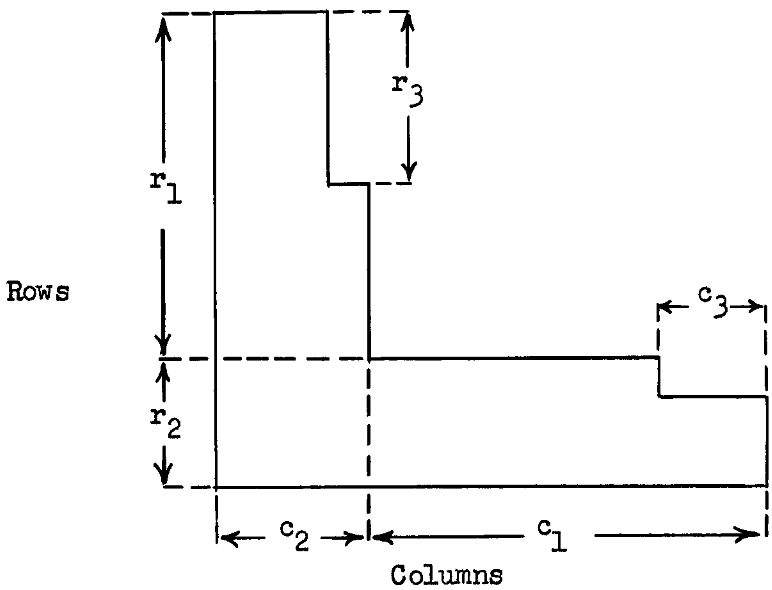

Figure 5. L-shaped design

8.2 Disjoint Rectangles Design

The second design investigated for simultaneously estimating σ_r^2 and σ_c^2 is designated as a "Disjoint Rectangles" design and is shown in Figure 6. This design consists of g rectangles, each containing r rows and c columns. The rectangles are disjoint in that each rectangle samples a different set of r rows and c columns. In this way several rows (gr) and columns (gc) are sampled.

The model for the disjoint rectangles design shown in Figure 6 is

$$y_{ijk\ell} = \mu + g_i + r_{ij} + c_{ik} + (rc)_{ijk} + e_{ijk\ell} \qquad (95)$$

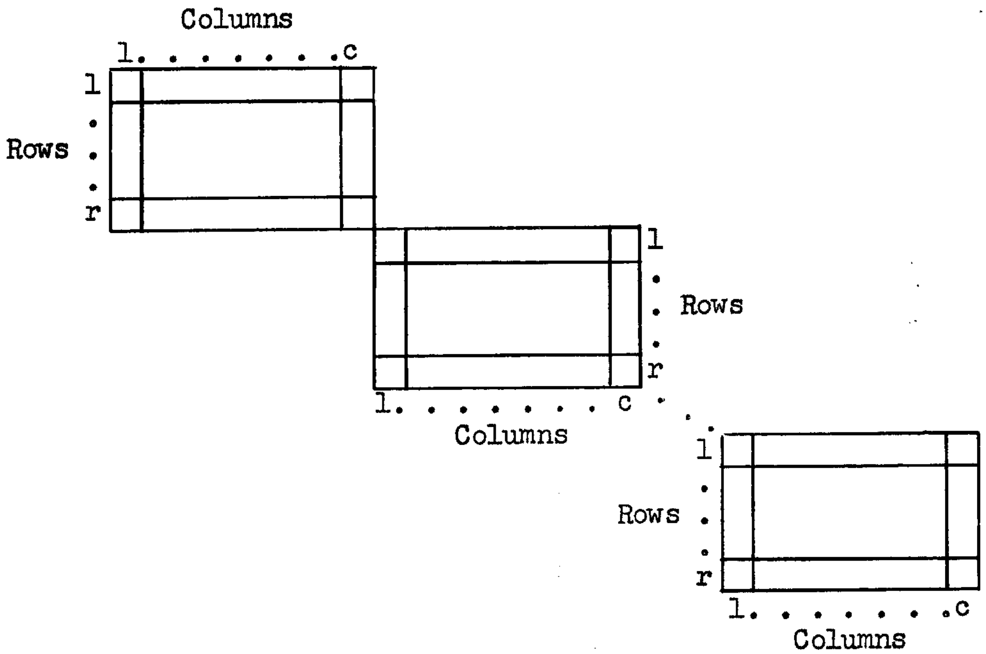

Figure 6. Disjoint rectangles design

where $i = 1,2,\ldots g$; $j = 1,2,\ldots r$; $k = 1,2,\ldots c$; $\ell = 1,2,\ldots n_{ijk}$; and the total number of observations is $\sum_{ijk} n_{ijk} = N$. $y_{ijk\ell}$ is the ℓ^{th} observation in the j^{th} row and k^{th} column of the i^{th} rectangle; μ is the general mean; the rectangle effects, g_i, are $NID(0,\sigma_g^2)$; the row effects within rectangles, r_{ij}, are $NID(0,\sigma_r^2)$; the column effects within rectangles, c_{ij}, are $NID(0,\sigma_c^2)$; the interaction effects within rectangles, $(rc)_{ijk}$, are $NID(0,\sigma_{rc}^2)$; the within cell effects, $e_{ijk\ell}$, are $NID(0,\sigma_e^2)$; and all effects are uncorrelated. The analysis of variance for the model given by (95) is shown in Table 10 where $n_{ijk} = n$ for all cells. If $\sigma_g^2 = 0$, G should be used to improve the estimates of σ_r^2 and σ_c^2. This case will not be considered in this dissertation.

Table 10. Analysis of variance for disjoint rectangles with n observations per cell

Source of variation	DF	MS	E(MS)
Rectangles	g-1	G	$\sigma_e^2 + n\sigma_{rc}^2 + nc\sigma_r^2 + nr\sigma_c^2 + nrc\sigma_g^2$
Rows	g(r-1)	R	$\sigma_e^2 + n\sigma_{rc}^2 + nc\sigma_r^2$
Columns	g(c-1)	C	$\sigma_e^2 + n\sigma_{rc}^2 + nr\sigma_c^2$
Interaction	g(r-1)(c-1)	I	$\sigma_e^2 + n\sigma_{rc}^2$
Error	grc(n-1)	E	σ_e^2
Total	N-1		

For a balanced case as shown in Table 10, Graybill and Wortham [1956] show that the estimator of the components of variance obtained by equating the mean squares to their expected values are the best unbiased quadratic estimators. The estimator for σ_r^2 is

$$\hat{\sigma}_r^2 = (R-I)/nc \ . \tag{96}$$

The variance of $\hat{\sigma}_r^2$ is

$$V(\hat{\sigma}_r^2) = \frac{2r}{N(r-1)nc}\left[(\sigma_e^2 + n\sigma_{rc}^2 + nc\sigma_r^2)^2 + \frac{(\sigma_e^2 + n\sigma_{rc}^2)^2}{c-1}\right] . \tag{97}$$

It will now be shown that (97) is minimized when n=1 if g is an integral multiple of n. Consider any disjoint rectangles design with the parameters: g, r, c, and $n_{ijk} = n$. From this design, a new design can be constructed with parameters: $g' = g/n$, $r' = nr$, $c' = nc$ and $n_{ijk} = 1$. The variance of $\hat{\sigma}_r^2$ for this design is from (97)

$$V'(\hat{\sigma}_r^2) = \frac{2r}{N(nr-1)c}\left[(\sigma_e^2 + \sigma_{rc}^2 + nc\sigma_r^2)^2 + \frac{(\sigma_e^2 + \sigma_{rc}^2)^2}{nc-1}\right]. \quad (98)$$

The term in [brackets] in (98) is obviously less than or equal to the similar term in (97). Since

$$\frac{1}{nr-1} \leq \frac{1}{(r-1)n} ,$$

$$\frac{2r}{N(nr-1)c} \leq \frac{2r}{N(r-1)nc} .$$

Hence, $V'(\hat{\sigma}_r^2) \leq V(\hat{\sigma}_r^2)$, where g is an integral multiple of n. Similarly, $V'(\sigma_c^2) \leq V(\sigma_c^2)$. Then, the best disjoint rectangles design with $n_{ijk} = 1$ should be found. A more general proof has not been attempted. It appears that in the case where $n_{ijk} = n$, setting $n=1$ is desirable.

The remaining discussion in this section will be limited to the case where $n=1$.

8.3 Comparison of L-shaped and Disjoint Rectangles Designs

In this thesis, these two designs will be compared for the special situation when $\sigma_r^2 = \sigma_c^2$ and $V(\hat{\sigma}_r^2) = V(\hat{\sigma}_c^2)$. Admittedly, this may be somewhat artificial, but a comparison on this basis should indicate the relative usefulness of these two designs in view of the fact that it has already been shown that allocation is not too sensitive to ρ. It is contemplated that further research will be devoted to the problem when the two components of variance are different.

Table 11 shows the results of some specific numerical examples, where $\rho = \sigma_r^2/(\sigma_e^2 + \sigma_{rc}^2) = \sigma_c^2/(\sigma_e^2 + \sigma_{rc}^2)$ and n=1. Since $\sigma_r^2 = \sigma_c^2$, it seemed reasonable to consider designs which give $V(\hat{\sigma}_r^2) = V(\hat{\sigma}_c^2)$. The value of the total number of observations, N^*, was allowed to vary slightly in order to obtain balance. In this case, comparisons were made on the basis of $N^*V(\hat{\sigma}_r^2)$. The reciprocal of this quantity is the amount of information per observation. In Table 11, $V = V(\hat{\sigma}_r^2)/2\sigma^4 = V(\hat{\sigma}_c^2)/2\sigma^4$. The design parameters are identified from Figures 5 and 6. The estimator of σ_r^2 for disjoint rectangles is given by (96) and the estimator for the L-shaped design is given by (28). The variances in the last column in Table 11, $N^*\hat{\hat{V}}^2$, are based on an estimator, $\hat{\hat{\sigma}}_r^2$, of σ_r^2 in which the last c_1 columns in Figure 5 are ignored. Otherwise, the estimator is the same as the one given by (28). The estimator, $\hat{\hat{\sigma}}_r^2$, could be improved slightly by pooling the error sum of squares from both parts of the design.

Several tentative conclusions can be drawn from Table 11. The best disjoint rectangles design is obtained when $r \approx 2 + 1/\rho$. When $\rho < 2$, the best disjoint rectangles design is better than the best L-shaped design. When $\rho \approx 2$, there is little difference between the best of each of the two designs. When $\rho > 2$, the best L-shaped design appears better than the best disjoint rectangles. As ρ becomes large, the estimator of $\hat{\hat{\sigma}}_r^2$ in the L-shaped design which discards the last c_1 columns becomes better than the estimator using all of the observations. This improvement is apparently a

Table 11. Comparison of disjoint rectangles and L-shaped designs with $V(\sigma_r^2) = V(\sigma_c^2)$, where $\sigma_r^2 = \sigma_c^2$

ρ	Disjoint rectangles				L-shaped					
	g	r=c	N^*	N^*V	r=c	$r_2=c_2$	$r_3=c_3$	N^*	N^*V	$N^*\hat{V}$
.5	4	3	36	3.38	10	2	0	36	4.24	5.00
.5	2	4	32	3.11	7	3	0	33	3.29	4.12
.5	1	6	36	3.24	6	6	0	36	3.24	3.24
1	9	2	36	10.00	10	2	0	36	10.89	10.00
1	4	3	36	8.25	7	3	0	33	9.46	10.08
1	2	4	32	8.44	6	6	0	36	9.84	9.84
2	9	2	36	26.0	14	2	8	36	-	25.6
2	4	3	36	24.8	10	2	0	36	34.8	26.0
2	2	4	32	27.1	7	3	0	33	31.6	30.2
4	9	2	36	82.0	16	2	12	36	-	79.8
4	4	3	36	84.8	10	2	0	36	125	82.0

result found earlier in Section 6.5 where it was shown desirable to keep the numbers of observations per row approximately equal. This was particularly true for large ρ. Also the L-shaped design which has partial rows and columns appears to be desirable when $\rho > 1$, where $c_2 = 2$. In this case a precise statement can be made concerning the use of partial columns (or rows). The analysis of variance ignoring the last c_1 columns is given in Table 12, where $c_2 = 2$. Let $r_4 = r - r_3$, the number of rows with two columns.

The estimator of σ_r^2 used is

$$\hat{\sigma}_r^2 = (R^* - E)/c_o \ .$$

In this case

$$\hat{\sigma}_r^2 = \frac{(r_4-1)R_1^* + (r-r_4)R_2^* - (r-1)E}{r + r_4 - 2} \ . \tag{99}$$

Table 12. Analysis of variance ignoring last c_1 columns

Source of variation	DF	MS	E(MS)
Columns	1	C	
Rows (adjusted for columns)	r-1	R^*	$\sigma^2 + c_o\sigma_r^2$
Component 1	r_4-1	R_1^*	$\sigma^2 + 2\sigma_r^2$
Component 2	r-r_4	R_2^*	$\sigma^2 + \sigma_r^2$
Error	r_4-1	E	σ^2
Total	N_1-1		

The variance of $\hat{\hat{\sigma}}_r^2$ is

$$V(\hat{\hat{\sigma}}_r^2) = \frac{2\sigma^4}{(r+r_4-2)^2}\left[(r_4-1)(1+2\rho_r)^2 + (r-r_4)(1+\rho_r)^2 + \frac{(r-1)^2}{r_4-1}\right], \quad (100)$$

where $\rho_r = \sigma_r^2/\sigma^2$.

To find the value of r_4 which minimizes (100), set the derivative with respect to r_4 equal to zero, subject to the restriction, $N_1 = r + r_4$. This leads to the following result. Use

$$r_4 = 1 + (N_1 - 2)/\rho_r\sqrt{2}\ , \qquad \rho_r \geq \sqrt{2}$$

$$r_4 = N_1/2\ , \qquad \rho_r \leq \sqrt{2}\ . \quad (101)$$

That is, when $\rho_r > \sqrt{2}$, one long column and one shorter (partial) column is better than using two columns of equal length. Of course, r_4 must be at least two in order to provide an estimator.

It will now be shown that an L-shaped design with $r_2 = c_2 = 2$ and $r = c$, where the last c_1 columns are neglected, is equivalent to the disjoint rectangles design with $r = c = 2$. For the disjoint rectangles, $g = N/4$. From (97), the variance of σ_r^2 is

$$V(\hat{\sigma}_r^2) = \frac{2\sigma^4}{N}\left[(1 + 2\rho_r)^2 + 1\right] . \qquad (102)$$

For the L-shaped design with two complete columns, $r_4 = r$. Since $r = c$ and

$$N = 2r + 2c - 4 \quad ,$$

$$r = \frac{N}{4} + 1 \quad .$$

Substituting these results into (100) gives

$$V(\hat{\hat{\sigma}}_r^2) = \frac{2\sigma^4}{N}\left[(1 + 2\rho_r)^2 + 1\right] . \qquad (103)$$

This is identical to (102) for disjoint rectangles. This result is borne out in Table 11. Similarly, the variances of $\hat{\sigma}_c^2$ would be identical. No restriction that $\sigma_r^2 = \sigma_c^2$ has been imposed.

It has been shown when $\rho_r > \sqrt{2}$ that the variance can be reduced for the L-shaped design by using a partial column and making the first column longer. In some instances, the L-shaped would be better. It appears, that a disjoint rectangles design with $r = c = 2$ would not be used. In Table 11 it is indicated that the best disjoint rectangles design has $r = c = 2$ for large ρ. Since the L-shaped design does better than 2x2 squares, this gives more evidence to support the use of the L-shaped design for large ρ.

In general, as ρ increases, the L-shaped design is favored over disjoint rectangles. Also, the analysis which discards the information in the other leg of the design in order to achieve balance appears more useful as ρ increases. When $\rho_r > \sqrt{2}$, the use of one long column and a second shorter column is advantageous. Similarly, for ρ_c. Disjoint rectangles appear better for small ρ. When ρ is approximately two, it does not appear that one design is particularly superior to the other. These observations must be taken with some reservation as the case for $\sigma_r^2 \neq \sigma_c^2$ was not investigated; also the size of N^* may have an influence on where one design is preferable to the other.

9.0 PROCEDURES TO ATTAIN SPECIFIED PRECISIONS OF ESTIMATES OF COMPONENTS OF VARIANCE

9.1 One-way Nested Classification

Consider the problem of finding the minimum value of N such that the variances of the estimators of the components of variance in the model given by (2) are less than or equal to specified values

$$V(\hat{\sigma}_a^2) \le p_1' \text{ and } V(\hat{\sigma}_e^2) \le p_2' \tag{104}$$

or

$$C.V.^2(\hat{\sigma}_a^2) \le p_1'\sigma_a^{-4} = p_1 \tag{105}$$

and

$$C.V.^2(\hat{\sigma}_e^2) \le p_2'\sigma_e^{-4} = p_2 \ . \tag{106}$$

The variance of $\hat{\sigma}_a^2$ is a complicated function, Crump [1954]. To find the minimum value of N which satisfies (105) would require a trial and error procedure. For a particular value of N' the optimal design is found by the procedure given by Crump. Then the variance of $\hat{\sigma}_a^2$ would be determined for the optimal design based on N'. If (105) was satisfied, then values of N less than N' would be tried. If (105) was not satisfied, then values of N greater than N' would be tried. This would continue until the smallest N satisfying (105) was found. Such a procedure would be extremely tedious. A simple scheme which gives the approximate minimum value of N which satisfies (104) will now be developed based on an approximation for the variance of $\hat{\sigma}_a^2$.

Consider the model given by (2) where the analysis of variance is given in Table 1. The estimator of σ_a^2 obtained by equating mean squares to their expected values is

$$\hat{\sigma}_a^2 = (A-W)/n_o \; . \tag{107}$$

Hammersley [1949] showed that the minimum variance of $\hat{\sigma}_a^2$ is obtained when $n_i = n$, i.e., equal numbers of observations per class. However, integer values can not always be obtained. If $N = ak + s$ where N, a, k, and s are integers, Crump [1954] showed the optimal design consists of s classes with (k + 1) observations and a - s classes with k observations each. That is, $n_i = k$ or $k + 1$ where k is an integer. In order to simplify further calculations, it will be shown for this case that $n_o \approx N/a$ for large N. The value of n_o is

$$n_o = \frac{1}{a-1}\left[N - \frac{\sum n_i^2}{N}\right] . \tag{108}$$

The case that deviates most from equal numbers per class where $n_i = k$ or $k + 1$ is when a/2 of the classes contain k observations and a/2 of the classes contain k + 1 observations. In this case

$$N = ak + a/2 \tag{109}$$

and (108) becomes

$$n_o = \frac{1}{a-1}\left[N - \frac{ak^2 + ak + a/2}{N}\right] = \frac{N}{a} - \frac{a}{4N(a-1)} \; .$$

Hence, as N increases

$$n_o \approx N/a \; . \tag{110}$$

The variance of $\hat{\sigma}_a^2$ is

$$V(\hat{\sigma}_a^2) = \left[V(A) + V(W)\right]/n_o^2 \tag{111}$$

where

$$V(W) = 2\sigma_e^4/(N-a) \ . \qquad (112)$$

Since the numbers of observations per class are nearly equal, $V(A)$ can be approximated by

$$V(A) \approx 2\sigma_e^4 \ (1 + n_o\rho)^2/(a-1) \qquad (113)$$

where $\rho = \sigma_a^2/\sigma_e^2$. Substituting from (112) and (113) into (111) gives

$$V(\hat{\sigma}_a^2) \approx \frac{2\sigma_e^4}{n_o^2}\left[\frac{(1 + n_o\rho)^2}{a-1} + \frac{1}{N-a}\right]. \qquad (114)$$

From (110), $a \approx N/n_o$, in order to minimize $V(\hat{\sigma}_a^2)$, giving

$$V(\hat{\sigma}_a^2) \approx \frac{2\sigma_e^4}{n_o}\left[\frac{(1 + n_o\rho)^2}{N-n_o} + \frac{1}{N(n_o-1)}\right]. \qquad (115)$$

Using the same allocation

$$V(\hat{\sigma}_e^2) \approx 2\sigma_e^4 n_o/N(n_o-1) \ . \qquad (116)$$

From (4), for large N and ρ not too small

$$n_o \approx (\rho + 1)/\rho \ . \qquad (117)$$

Substituting this result into (115) gives for large N

$$NV(\hat{\sigma}_a^2) \approx \frac{2\sigma_e^4\rho^2}{(\rho+1)}\left[\frac{N\rho^2 + 5N\rho + 4N - \rho - 1}{N\rho - \rho - 1}\right]. \qquad (118)$$

As N increases

$$NV(\sigma_a^2) \longrightarrow 2\sigma_e^4\rho(\rho + 4) \qquad (119)$$

for the approximate minimum obtainable value of $NV(\hat{\sigma}_a^2)$ for large N.

From (119) the approximate minimum value for the square of the coefficient of variation of $\hat{\sigma}_a^2$ is given by

$$N{\cdot}C.V.^2(\hat{\sigma}_a^2) \approx 2(\rho + 4)/\rho. \qquad (120)$$

Then $C.V.^2(\hat{\sigma}_a^2) \lesssim p_1$ when

$$2(\rho + 4)/N\rho \leq p_1$$

or when

$$N \geq 2(\rho + 4)/\rho p_1 = N_a \; . \qquad (121)$$

Table 13 shows the approximate values of N required to satisfy a specified $C.V.(\hat{\sigma}_a^2)$ for various values of ρ. These results are quite good even for small values of N. For example, C.V. is .53 for ρ=2, N=24 and ρ=4, N=16.

Table 13. Approximate total number of observations (N_a) required for estimation of σ_a^2

ρ	Specified $C.V.(\hat{\sigma}_a^2)$ .10	.20	.30	.50
1/4	3400	850	378	136
1/2	1800	450	200	72
1	1000	250	112	40
2	600	150	67	24
4	400	100	45	16

Substituting from (117) into (116) gives the approximate variance of $\hat{\sigma}_e^2$ from the optimal design for estimating σ_a^2 for large N

$$NV(\hat{\sigma}_e^2) \approx 2\sigma_e^4(\rho + 1) \; . \qquad (122)$$

The square of the coefficient of variation of $\hat{\sigma}_e^2$ is

$$C.V.^2(\hat{\sigma}_e^2) \approx 2(\rho + 1)/N \; . \qquad (123)$$

Now, $C.V.^2(\hat{\sigma}_e^2) \lesssim p_2$ when

$$2(\rho + 1)/N \leq p_2,$$

or when

$$N \geq 2(\rho + 1)/p_2 = N_e \ . \qquad (124)$$

When $N_e \leq N_a$, then (105) and (106) are approximately satisfied for large N by choosing $N = N_a$ and using the optimal design for estimating σ_a^2.

When $N_e > N_a$, then additional observations are required so that $C.V.^2(\hat{\sigma}_e^2) \lesssim p_2$. The approximate number of classes required for $C.V.^2(\hat{\sigma}_a^2) \leq p_1$ is $a \approx N/n_o$. If observations are added to the classes already sampled, then each observation adds one degree of freedom to the estimate of σ_e^2. From (112), it follows that for a fixed a, the allocation of the N samples to the a classes has no affect on $V(\hat{\sigma}_e^2)$. But, these extra observations may as well be added so that the number of observations per class are almost equal in order to minimize $V(\hat{\sigma}_a^2)$. The number of additional observations required to satisfy (106) when $N_e > N_a$ is approximately $N_e - N_a$ for large N.

9.2 Two-way Crossed Classification

Consider a procedure such that the variances of the estimates of the row and column components of variance of the model given by (1) are less than or equal to specified values

$$V(\hat{\sigma}_r^2) \leq p_1' \qquad (125)$$

and

$$V(\hat{\sigma}_c^2) \leq p_2' \qquad (126)$$

or

$$\text{C.V.}^2(\hat{\sigma}_r^2) \leq p_1'\sigma_r^{-4} = p_1 \qquad (127)$$

and

$$\text{C.V.}^2(\hat{\sigma}_c^2) \leq p_2'\sigma_c^{-4} = p_2 \,. \qquad (128)$$

Consider the L-shaped design discussed in Section 8.1. Substituting from (93) into (56) gives the approximate minimum value obtainable for $V(\hat{\hat{\sigma}}_r^2)$, for large N_1,

$$\tilde{V}(\hat{\hat{\sigma}}_r^2) = 2\sigma^4\rho_r(\rho_r + 4)/N_1 \,.$$

Similarly,

$$\tilde{V}(\hat{\hat{\sigma}}_c^2) = 2\sigma^4\rho_c(\rho_c + 4)/N_2 \,.$$

Or,

$$\widetilde{\text{C.V.}}^2(\hat{\hat{\sigma}}_r^2) = 2\ (\rho_r + 4)/N_1\rho_r \qquad (129)$$

and

$$\widetilde{\text{C.V.}}^2(\hat{\hat{\sigma}}_c^2) = 2\ (\rho_c + 4)/N_2\rho_c \,. \qquad (130)$$

Then (127) is approximately true when

$$2\ (\rho_r + 4)/N_1\rho_r \leq p_1$$

or

$$N_1 \geq 2\ (\rho_r + 4)/\rho_r p_1 \,. \qquad (131)$$

Formula (128) is approximately true when

$$2\ (\rho_c + 4)/N_2\rho_c \leq p_2$$

or

$$N_2 \geq 2\ (\rho_c + 4)/\rho_c p_2 \,. \qquad (132)$$

Thus, the total number of observations required is

$$N \approx N_1 + N_2 - (\rho_r + 1)(\rho_c + 1)/\rho_r\rho_c$$

or from (131) and (132)

$$N \approx 2(\rho_r + 4)/\rho_r p_1 + 2(\rho_c + 4)/\rho_c p_2 - (\rho_r + 1)(\rho_c + 1)/\rho_r \rho_c \, . \quad (133)$$

Consider the disjoint rectangles design described in Section 8.2. The variance of $\hat{\sigma}_r^2$ is given by (97). The coefficient of variation squared of $\hat{\sigma}_r^2$ is less than p_1 when

$$C.V.^2(\hat{\sigma}_r^2) = \frac{2}{g(r-1)c^2\rho_r^2}\left[(1 + c\rho_r)^2 + \frac{1}{c-1}\right] \leq p_1 \, . \quad (134)$$

Similarly,

$$C.V.^2(\hat{\sigma}_c^2) = \frac{2}{g(c-1)r^2\rho_c^2}\left[(1 + r\rho_c)^2 + \frac{1}{r-1}\right] \leq p_2 \, . \quad (135)$$

A procedure which leads to the minimum value of N such that both (134) and (135) are satisfied is as follows. Compute for any particular values of $r \geq 2$ and $c \geq 2$:

$$g_r = \frac{2}{p_1(r-1)c^2\rho_r^2}\left[(1 + c\rho_r)^2 + \frac{1}{c-1}\right] \quad (136)$$

and

$$g_c = \frac{2}{p_2(c-1)r^2\rho_c^2}\left[(1 + r\rho_c)^2 + \frac{1}{r-1}\right] \, . \quad (137)$$

Choose g as the smallest integer greater than or equal to the maximum $[g_r, g_c]$. Then, $N = grc$. Repeat this procedure for various combinations of r and c until the minimum value of N is found. In this way, both (134) and (135) are satisfied for a minimum value of N.

10.0 SOME FURTHER RESULTS FOR A ONE-WAY NESTED CLASSIFICATION

Consider the one-way nested classification described by the model given in (2). Suppose it is desired to estimate the total variance, $V(y_{ij}) = \sigma_a^2 + \sigma_e^2$, with a fixed cost. Consider the problem of minimizing $V(\sigma_a^2 \hat{+} \sigma_e^2)$ with a fixed cost.

For a given value of the number of classes, a, the degrees of freedom for error is fixed, N-a. Thus, the variance of $\hat{\sigma}_e^2$ is fixed since $V(\hat{\sigma}_e^2) = 2\sigma_e^4/(N-a)$. The variance of $\hat{\sigma}_a^2$ is minimized for a given a by making the numbers of observations per class as nearly equal as possible, Crump [1954].

From the expected mean squares in Table 1 an estimator of $(\sigma_a^2 + \sigma_e^2)$ is obtained by

$$\sigma_a^2 \hat{+} \sigma_e^2 = \left[A + (n_o-1)W\right]/n_o \;. \tag{138}$$

The variance of this estimator of $\sigma_a^2 + \sigma_e^2$ is

$$V(\sigma_a^2 \hat{+} \sigma_e^2) = \left[V(A) + (n_o-1)^2 V(W)\right]/n_o^2 \tag{139}$$

where

$$V(W) = 2\sigma_e^4/(N-a) \tag{140}$$

and as seen previously

$$V(A) \approx 2(\sigma_e^2 + n_o\sigma_a^2)^2/(a-1) \tag{141}$$

since the numbers of observations per class are almost equal. Substituting from (140) and (141) into (139) gives

$$V(\sigma_a^2 \hat{+} \sigma_e^2) \approx \frac{2\sigma_e^4}{n_o^2}\left[\frac{(1 + n_o\rho)^2}{a-1} + \frac{(n_o-1)^2}{N-a}\right]. \tag{142}$$

From (110), $N \approx n_o a$. Substituting this result into (142) gives for large N

$$V(\sigma_a^2 \hat{+} \sigma_e^2) \approx \frac{2\sigma_e^4}{n_o^2} \left[\frac{2an_o\rho + an_o^2\rho^2 + an_o - n_o + 1}{a(a-1)} \right] . \qquad (143)$$

Since $an_o \approx N$, the numerator in the [brackets] of (143) is approximately for large N

$$2an_o\rho + an_o^2\rho^2 + an_o .$$

Substituting this result into (143) gives

$$V(\sigma_a^2 \hat{+} \sigma_e^2) \approx 2\sigma_e^4 (2\rho + n_o\rho^2 + 1)/n_o(a-1) . \qquad (144)$$

Let

c_a = cost of sampling a class

c_e = cost of obtaining an observation in a class.

The total fixed cost, C, is

$$C = ac_a + Nc_e \approx ac_a + an_oc_e . \qquad (145)$$

Hence,

$$a \approx C/(c_a + n_oc_e) . \qquad (146)$$

Differentiating (144) with respect to a, subject to the condition given by (146), and setting this result equal to zero gives a quadratic equation in a. Denote the solution of this equation by a_1. As a first approximation to the value of a which minimizes the variance of $(\sigma_a^2 \hat{+} \sigma_e^2)$, choose the integer closest to a_1 and N the largest integer less than or equal to $(C - ac_a)/c_e$. Allot the N observations such that the numbers of observations per class are as nearly equal as possible.

11.0 SUMMARY AND CONCLUSIONS

11.1 The Problem

A sample of size N is to be selected from a population which can be classified in two ways, called rows and columns. In selecting the sample, it is assumed that r rows and c columns have been randomly selected from the large number available and that for each row and column cell (i, j) a certain number of measurements or observations (n_{ij}) are obtained at random. The measured value (y) of a given sample taken from the i^{th} row and j^{th} column is assumed to be represented by the following linear random model:

$$y_{ijk} = \mu + r_i + c_j + (rc)_{ij} + e_{ijk} ,$$

where μ is the average value of the measurements in the population and k represents the particular sample selected. The random effects, r_i, c_j, $(rc)_{ij}$, and e_{ijk} represent row, column, interaction, and sampling (within a row-column cell) variation. It is assumed that their effects are normally and independently distributed with zero means and variances, σ_r^2, σ_c^2, σ_{rc}^2, and σ_e^2, respectively.

The purpose of this investigation was to devise the best method of selecting a sample of size N so as to obtain minimum variance estimators of the parameters μ, σ_r^2, σ_c^2, σ_{rc}^2 and σ_e^2.

11.2 Optimal Allocation

If we use as the estimator of μ, $\hat{\mu} = \sum\sum\sum y_{ijk}/N$, where the summation is over the members of the sample, the variance of $\hat{\mu}$ is minimized if each observation is taken from a different row and column, so that $r = c = N$. After a permutation of rows and columns, $n_{ij} = 1$ for $i=j$ and $n_{ij} = 0$ for $i \neq j$ $(i,j = 1,2,\ldots, N)$.

If we consider any linear function of the components in y_{ijk}, the variance of the linear function being designated as σ^2, the lower bound for the variance of an unbiased quadratic estimator of σ^2 is $2\sigma^4/(N-1)$. Although it is not always possible to attain this lower bound, estimators of the following can attain the lower bound:

(1) σ_e^2, if all N observations are in one cell, $r = c = 1$ and $n_{11} = N$;

(2) $\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2$, if only one column is selected with N rows, with one observation per cell, $r = N$, $c = 1$ and $n_{i1} = 1$ $(i = 1,2,\ldots, N)$;

(3) $\sigma_e^2 + \sigma_{rc}^2 + \sigma_c^2$, if rows and columns are reversed in (2);

(4) the total variance, $\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2 + \sigma_c^2$, if each observation is selected from a different row and column, so that $r = c = N$, $n_{ij} = 1$ for $i=j$ and $n_{ij} = 0$ for $i \neq j$.

For other linear forms, in particular for σ_r^2 and σ_c^2, it is not possible to attain the lower bound. The variance of the estimator is materially affected by the sampling procedure. Also, optimal allocation will vary from one estimator to another.

The estimator used for σ_r^2 was obtained by equating mean squares to their expected values in an analysis of variance based on the method of fitting constants, where the row mean square is adjusted for both row and column effects. With unequal numbers of observations per cell, the variance of $\hat{\sigma}_r^2$ is a very complicated function.

This investigation was limited to designs of the type where there might be some vacant cells and each occupied cell contained n observations, so that $n_{ij} = 0$ or n. These designs led to a simple estimator of σ_r^2 based on the difference of two mean squares. It was proved for this type of design that the variance of $\hat{\sigma}_r^2$ is minimized when $n_{ij} = 0$ or 1 (n=1). For a given number of rows, r, the optimal design with $n_{ij} = 0$ or 1 consists of c-1 columns and all r rows in common and one column with u ($0 \leq u \leq s \leq r$, $u \neq 1$) of the r rows, where

$$N = r(c-1) + s .$$

An approximate variance of $\hat{\sigma}_r^2$ was developed to obtain a first approximation, $\tilde{r}$, for the number of rows which minimizes the exact variance of $\hat{\sigma}_r^2$. The value of $\tilde{r}$ is given by

$$\tilde{r} = (N - \tilde{c} + \tilde{c}_o)/\tilde{c}_o ,$$

in which $\tilde{c}$ is the smallest integer greater than or equal to $\tilde{c}_o$ and

$$\tilde{c}_o = \frac{\rho_r(N - 1/2) + (N - 1/2) + 1}{\rho_r(N - 1/2) + 2} ,$$

where $\rho_r = \sigma_r^2/(\sigma_e^2 + \sigma_{rc}^2)$. Integers below $\tilde{r}$ and integers above $\tilde{r}$ are tried until the value of r is obtained which minimizes the exact variance of $\hat{\sigma}_r^2$. The exact variance of $\hat{\sigma}_r^2$ for this form of the optimal design is

$$V(\hat{\sigma}_r^2) = \frac{2\sigma^4}{(N-c)^2}\left[\begin{array}{l}(r-t)\left[1 +(c-1)\rho\right]^2 + (t-1)(1+c\rho)^2 \\ + (r-1)^2/ (N-r-c+1)\end{array}\right],$$

where $\sigma^2 = (\sigma_e^2 + \sigma_{rc}^2)$, t=u for $u \geq 2$ and t=1 for u=0. For each value of r tried, the value of u must first be determined to minimize the variance of $\hat{\sigma}_r^2$. Generally, the best value of u is 0 or s; when s=r, u=r.

The determination of the optimal design depends on the value of ρ_r. It was shown by several numerical examples that if the value of ρ_r actually used in constructing the optimal design differed from the true value by a factor of two, the relative efficiency of the design is still generally greater than 90%.

When $\sigma_{rc}^2 = 0$, it was shown that the optimal design for estimating σ_r^2 consists of one column. Then, the analysis of variance reduces to a one-way nested classification for which the optimal allocation for estimating σ_r^2 is given by Crump [1954].

From the same analysis of variance used to estimate σ_r^2 with n_{ij} = 0 or 1, a simple estimator of ρ_r is obtained from the ratio of the row to interaction mean squares. The form of the design and procedures which lead to the design which minimizes the variance of $\hat{\rho}_r$ are analogous to those obtained for estimating σ_r^2 except that the first trial value for r is determined by using

$$\tilde{c}_o = \frac{2\rho_r(N - 1/2) + (N - 1/2) + 1}{\rho_r(N - 1/2) + \rho_r + 2}$$

and the variance of $\hat{\rho}_r$ for the form of the optimal design is

$$V(\hat{\rho}_r) = \frac{2\ (r-1)^{-1}}{(N-c)^2(N-r-c-3)} \left[\begin{array}{l} (N-r-c-1)(t-1)(r-t)\rho_r^2 \\ + (N-c-2)(r-t + (N-c)\rho_r)^2 \end{array} \right]$$

where t=u for $u \geq 2$, t=1 for u=0.

The determination of the optimal design for estimating ρ_r requires some previous knowledge of the value of ρ_r, which is the quantity one is attempting to estimate. Even if the value of ρ_r actually used to determine the optimal design differs from the true value by a factor of four, several numerical examples showed that the relative efficiency of the design is still generally greater than 80%.

When $\sigma_{rc}^2 = 0$, the optimal design for estimating ρ_r then was shown to consist of one column (j=1) with no restriction on the n_{i1}'s, i.e. $n_{ij} \geq 0$. The analysis of variance with one column reduces to a one-way nested classification for which the optimal allocation for estimating ρ_r is given by Crump [1954].

All of the results for estimating σ_r^2 and ρ_r apply to the estimation of σ_c^2 and ρ_c simply by interchanging r and c.

A few tentative conclusions were obtained concerning the simultaneous estimation of σ_r^2 and σ_c^2. The investigation was limited to two types of designs: (1) an L-shaped design consisting of N_1 observations with several rows and a few columns and N_2 observations with several columns and few rows and (2) a disjoint rectangles design consisting of several non-overlapping rectangles, each with r rows and c columns. The comparison of these two designs was limited primarily to some numerical examples. In order to simplify the comparison, the only cases considered were for $\sigma_r^2 = \sigma_c^2$ so that $\rho_r = \rho_c = \rho$. This may be somewhat artificial, but it should lead to some general notions. The designs were constructed so that $V(\hat{\sigma}_r^2) = V(\hat{\sigma}_c^2)$.

As ρ increases, the usefulness of the L-shaped design increases. It was shown that the N_1 observations should contain two or more columns when $\rho \leq \sqrt{2}$ and one long column and one shorter column when $\rho > \sqrt{2}$. Also, as ρ increases, the variance of $\hat{\sigma}_r^2$ is smaller when only the N_1 observations are used. The gain from discarding observations is due to the near balance which is achieved in this manner.

The usefulness of the disjoint rectangles design increases as ρ decreases, r and c increase for the optimal disjoint rectangles. When ρ becomes large, the optimal values of r and c equal two. In this case, it was shown that an L-shaped design could be constructed which is better than disjoint rectangles. This also supports the observation that L-shaped designs are better for large ρ. The optimal designs of each type appear to be about equivalent when ρ is around the value of two.

A procedure was developed to find the approximate minimum value of N required in order that both $V(\hat{\sigma}_a^2)$ and $V(\hat{\sigma}_e^2)$ are less than specified values in a one-way nested classification. The approximate minimum values of N required in order that both $V(\hat{\sigma}_r^2)$ and $V(\hat{\sigma}_c^2)$ are less than specified values were found for the L-shaped design and a procedure was developed for finding the exact minimum value of N for a disjoint rectangles design.

A procedure was developed for a one-way nested classification which minimizes the approximate variance of the estimator of the total variance, $\sigma_a^2 + \sigma_e^2$, subject to a fixed total cost, where the cost of sampling a class is not necessarily equal to the cost of sampling within a class.

11.3 Suggested Future Research

(1) Construction of more general designs without the restriction that $n_{ij} = 0$ or n. Apparently this will be very difficult because of the complicated general variance function.

(2) Construction of designs to estimate σ^2_{rc}, σ^2_{rc}/σ^2_e, σ^2_r/σ^2_e, and σ^2_c/σ^2_e. These require that some n_{ij}'s must exceed one.

(3) Study of robustness of the designs to non-normality and to inequality of some of the components. For example, estimate σ^2_r when σ^2_e and σ^2_{rc} vary from column to column.

(4) Development of procedures to combine estimators when the analysis of variance contains more mean squares than components of variance to be estimated.

(5) Consideration of estimators other than those based on the usual analysis of variance. Some simple examples indicate that this should receive serious consideration.

(6) Evaluation of designs on the basis of information per observation rather than on the variance of the estimator for fixed N, or on the basis of other cost considerations.

(7) Development of a sequential estimation procedure, with at least two stages, the first stage design to be used to estimate values of ρ to be used for designing the second stage.

(8) Extension of the results to multi-stage nested and crossed classifications, with both fixed and random components.

(9) Study of the effect of the n_{ij} being random, due to natural or other causes.

(10) Consider other designs and estimating procedures when several variance components must be estimated.

LIST OF REFERENCES

Anderson, R. L. and Bancroft, T. A., 1952. Statistical Theory in Research. McGraw-Hill Book Co., Inc. New York.

Baines, A. H. J., 1944. On the economical design of statistical experiments. Ministry of Supply Advisory Service on Statistical Methods and Quality Control. Tech. Rep. No. Q.C./R/15.

Crump, P. P., 1954. Optimal designs to estimate the parameters of a variance component model. Unpublished Ph.D. Thesis, North Carolina State College, Raleigh.

Graybill, F. A. and Wortham, A. W., 1956. A note on uniformly best unbiased estimators for variance components. J. Amer. Stat. Assoc. 51, 266-268.

Hammersley, J. M., 1949. The unbiased estimate and standard error of the interclass variance. Metron 15, 189-205.

Kempthorne, O., 1952. The Design and Analysis of Experiments. John Wiley and Sons, Inc. New York.

Lancaster, H. O., 1954. Traces and cumulants of quadratic forms in normal variables. J. Roy. Stat. Soc. (B) 16, 247-254.

Shah, B. V., 1959. A generalization of partially balanced incomplete block designs. Annals Math. Stat. 30, 1041-1050.

Whittle, P., 1953. The analysis of multiple stationary time series. J. Roy. Stat. Soc. (B) 15, 125-139.

ABSTRACT

GAYLOR, DAVID WILLIAM. The Construction and Evaluation of Some Designs for the Estimation of Parameters in Random Models. (Under the direction of RICHARD LOREE ANDERSON).

This dissertation considers methods of sampling, with fixed sample size N, which would lead to good estimates of components of variance in a two-way crossed classification model with n_{ij} observations in the (i,j) cell:

$$y_{ijk} = \mu + r_i + c_j + (rc)_{ij} + e_{ijk} ,$$

where the effects r_i, c_j, $(rc)_{ij}$, and e_{ijk} are normally and independently distributed random variables with zero means and variances σ_r^2, σ_c^2, σ_{rc}^2, and σ_e^2, respectively.

It was shown that the lower bound for the variance of an unbiased quadratic estimator of a linear function of components of variance with expected value, σ^2, is $2\sigma^4/(N-1)$ where N is the total number of observations. Procedures which achieved the lower bounds were determined for estimating σ_e^2, $(\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2)$, $(\sigma_e^2 + \sigma_{rc}^2 + \sigma_c^2)$, and $(\sigma_e^2 + \sigma_{rc}^2 + \sigma_r^2 + \sigma_c^2)$. In estimating other functions of the variance components, optimal allocation depends upon the estimators used. The estimator for σ_r^2 was obtained by equating mean squares to their expected values in an analysis of variance based on the method of fitting constants, where the row mean squares is adjusted for column effects and interaction is adjusted for both row and column effects.

A procedure was developed to minimize the variance of $\hat{\sigma}_r^2$ when $n_{ij} = 0$ or n; for this case, it was shown that n = 1.

From the same analysis of variance used to estimate σ_r^2, an estimator of $\sigma_r^2/(\sigma_e^2 + \sigma_{rc}^2)$ can be obtained from the ratio of the row to interaction mean squares. The form of the design of the type $n_{ij} = 0$ or 1 which minimized the variance of this estimator was determined.

Allocation for estimating σ_c^2 and $\sigma_c^2/(\sigma_e^2 + \sigma_{rc}^2)$ were treated similarly.

A few tentative conclusions on the simultaneous estimation of σ_r^2 and σ_c^2 were obtained. Two types of designs were compared.

Procedures were developed for both the one-way nested and two-way crossed classifications to find the approximate minimum value of N such that the variance of the estimators of components of variance were less than specified values.

A procedure was developed for the one-way nested classification which minimizes the approximate variance of an estimator of the total variance subject to a fixed total cost where the cost of sampling classes is not necessarily equal to the cost of sampling within classes.