# THE ESTIMATION OF A MISSING VALUE IN A LATTICE DESIGN USING INTER- AND INTRA- BLOCK INFORMATION

Submitted as a

MAJOR THESIS

to the

FACULTY OF THE

VIRGINIA POLYTECHNIC INSTITUTE

In Partial Fulfillment of the Requirements

for the

DEGREE OF MASTER OF SCIENCE IN STATISTICS

by

JAMES HENRY DAVIDSON

March, 1945

APPROVED

In charge of Statistics

Head, Agricultural Economics, Rural Sociology and Statistics

Dean of Agriculture

# THE ESTIMATION OF A MISSING VALUE IN A LATTICE DESIGN USING INTER- AND INTRA-BLOCK INFORMATION.

## ACKNOWLEDGEMENTS

When this work was contemplated, the end was certainly not in sight and no one knew just where the answer might have obscured itself. The existing literature, though background, formed a rather uncertain spring board from which to approach the problem; so that many failing starts were unavoidable before the actual course was taken. For this reason, I cannot be too grateful to Dr.Boyd Harshbarger for his putting at my disposal, his recent researches on the analysis of the various lattice designs. Without his numerous constructive criticisms, his suggestions, and his gracious interest, this work could never have been possible.

My sincere thanks go also, to Mrs. Rembold for not only preparing a rough draft from the manuscript, a particularly difficult job considering the type of study, but also for preparing the final copies for presentation.

Blacksburg, Virginia
February, 1945

CONTENTS

# INTRODUCTION

In the field of statistical research and analysis, particularly in agriculture, it often becomes desirable to test a large number of varieties or other factors at one time. Any experimental design such as the Latin square, which uses as many replications as there are varieties, is clearly too bulky and expensive to be of any help. The system of randomized blocks is also unsatisfactory, because each block must contain too many plots to be efficient. The use of control plots to furnish an index of block variation is also inefficient, because the precision is reduced, due to the fact that more plots are allotted to the control than to any other variety.

In 1936 Yates[1] devised the lattice square or psuedo-factorial type of arrangement. It has the advantage of being able to compare a large number of varieties and at the same time to eliminate variations due to blocks of values. The latter operation is accomplished by rearranging the order of varieties in the replications by changing the system of allotting varieties to blocks, as will be shown in the next section. The arrangement may be done in a number of ways, but the most popular designs are those containing three arrangements or the triple lattice, and those containing two arrangments, or the lattice.

All too often in experiments carried out in practice, a value is either very questionable compared with the other values, or it is missing altogether. In such a case, the symmetry of the experiment and the orthogonal method of analysis breaks down. The first recognition of this problem was made by Wishart and Allen[2] in 1930, when they derived formulae for use with the randomized blocks and Latin square. Their work was later amplified by Yates[3] and still later by Cornish[4]. However, Cornish was the only one to attempt the problem for the lattice designs. In his latest paper, various formulae are developed using only information obtained within blocks for his estimate. In fact, the following statement is made in his summary, "The determination of estimates which will enable interblock information to be recovered presents a different problem which is now under consideration."

This paper attempts to present an adequate formula for the estimation of a missing block, using interblock as well as intrablock information. The value obtained will be shown to give a closer approximation to what the true value would have been, depending on the amount of interblock variation. Also it follows that since a closer approach is made to the missing plot, a more accurate estimate is made of all standard errors and variances involved. An actual example will be worked to illustrate the new method of estimation.

## ANALYSIS OF LATTICE SQUARE DESIGN

The general construction and analysis of lattice designs have been formally given by Yates[5], Cochran[6], and Harshbarger[7]; hence, the symbols used here will coincide as nearly as possible to those already embodied in the nomenclature. A general discussion of the lattice and its theory of analysis will precede the numerical problem selected for illustration.

The number of treatments or varieties tested by the lattice is a perfect square, $p^2$, such that each square or incomplete block will have p varieties on a side. The number of exact replications in each group will be denoted by q, and the two different types of arrangements used will be designated by Group X and Group Y. Using subscript numerals to signify varieties, the two groups are illustrated in Figure I.

Group X

| | | | | total | block |
|---|---|---|---|---|---|
| $v_1$ | $v_2$ | $v_3$ | $\cdots v_p$ | $X_{1.}$ | (1) |
| $v_{p+1}$ | $v_{p+2}$ | $v_{p+3}$ | $\cdots v_{2p}$ | $X_{2.}$ | (2) |
| . | . | . | . | . | (3) |
| . | . | . | . | . | |
| . | . | . | . | . | |
| $v_{p^2-p+1}$ | $v_{p^2-p+2}$ | $v_{p^2-p+3}$ | $\cdots v_{p^2}$ | $X_{p.}$ | (p) |
| $X_{.1}$ | $X_{.2}$ | $X_{.3}$ | $X_{.p}$ | $X_{..}$ | |

Group Y

| | | | total |
|---|---|---|---|
| $v_1$ $v_{p+1}$ | . . | $v_{p^2-p+1}$ | $Y_{1.}$ |
| $v_2$ $v_{p+2}$ | . . | $v_{p^2-p+2}$ | $Y_{2.}$ |
| $v_3$ $v_{p+3}$ | . . | $v_{p^2-p+3}$ | $Y_{3.}$ |
| . . | | . | . |
| . . | | . | . |
| $v_p$ $v_{2p}$ | | $v_{p^2}$ | $Y_{p.}$ |
| $Y_{.1}$ $Y_{.2}$ | | $Y_{.p}$ | $Y_{..}$ |

Figure I

The row and column totals are indicated at the sides and bottom of each square respectively, with the grand total at the lower right corner of each group. The values of Figure I may be considered to represent the totals of all the q like replications for each group.

In practice, the blocks as well as the varieties within each block, are completely randomized, usually with the help of tables of random numbers. Also, each of the q replications for each group is randomized separately so that an unbiased estimate of each variety or treatment is available.

To make an analysis of the design in Figure I, let it be assumed that all the various effects operate additively. Then let the yield $Y_{eij}$ be located as the jth variety in the ith block in the eth replicate, and as yet irrespective of group. Then, the following equation can be formed:

$$Y_{eij} = \mu + \alpha_e + \beta_{ie} + V_{je} + \epsilon_{eij} \qquad (1)$$

where $\mu$ is the true population mean, $\alpha_e$ is the effect of the eth replication, $\beta_{ie}$ is the effect of the ith block in the eth replicate, $V_{je}$ is the population effect of the jth variety still in the eth replicate, and $\epsilon_{eij}$ is the experimental error of the eijth plot which is allocated to chance. Further, it is assumed that the experimental errors per plot are independent and normally distributed, and that the sum of the effects for replications, blocks, and varieties, is zero.

Using the above restrictions, the effects may be estimated by the method of least squares. Thus it is necessary to minimize:

$$S(Y_{eij} - mz_1 - r_e z_{2e} - \frac{(b_{ei} - b'_{ei})}{2} z_{3ei} - \frac{(b_{ei} + b'_{ei})}{2} z_{4ei} - vz_{5j})^2 + \lambda_1 \sum_{e=1}^{2q} r_e \qquad (2)$$

$$+ \sum_{e=1}^{2q} \lambda_{2e} \sum_{i=1}^{p} \frac{(b_{ei} - b'_{ei})}{2} + \sum_{e=1}^{2q} \lambda_{3e} \sum_{i=1}^{p} \frac{(b_{ei} + b'_{ei})}{2} + \ldots + \lambda_n \sum_{j=1}^{p^2} v_j = \pi$$

where $z_1 \ldots\ldots z_{5j}$ are the independent multipliers which will be used as Kronecker deltas, and the values of $Y_{eij}$ are dependent on them. The primes indicate corresponding values in the exact replications. For the lattice design, $z_1 = 1$ for all cases; $z_2 = 1$ for all values of plots in replicate e, but = 0 elsewhere; $z_3 = 1$ for all values of plots in the ith block, and = -1 in the exact replicate of the ith block, but = 0 elsewhere; $z_4 = 1$ for the entire ith block as well as for its exact replicate, but = 0 elsewhere; $z_5 = 1$ for the jth variety, but = 0 elsewhere.

By employing these restricted values of the z's and by virtue of the fact that all the expression to the right of the squared parenthesis drops out when partial derivatives are taken, a series of simple, normal equations

may be formed to express the various effects. Thus on partial differentiation of (2) with respect to the grand mean, m, one obtains:

$$2qp^2m + p^2 \sum_{1}^{2q} r_e + 2p \sum_{e=1}^{q} \sum_{i=1}^{p} \frac{(b_{ei} - b'_{ei})}{2} + 2p \sum_{e=1}^{q} \sum_{i=1}^{p} \frac{(b_{ei} + b'_{ei})}{2}$$

$$+ 2q \sum_{j=1}^{p^2} v_j = X.. + Y.. = T.. \qquad (3)$$

when:

$$\frac{\partial \pi}{\partial m} = 0$$

Likewise differentiating with respect to replication ($r_1$), one gets:

$$p^2m + p^2r_1 + p \sum_{e=1}^{q} \sum_{i=1}^{p} \frac{(b_{ei} - b'_{ei})}{2} + p \sum_{e=1}^{q} \sum_{i=1}^{p} \frac{(b_{ei} + b'_{ei})}{2} + \sum_{j=1}^{p^2} v_j = R_1$$

when:

$$\frac{\partial \pi}{\partial r_1} = 0$$

or generally for the eth replicate:

$$p^2m + p^2r_e + p \sum_{e=1}^{q} \sum_{i=1}^{p} \frac{(b_{ei} - b'_{ei})}{2} + p \sum_{e=1}^{q} \sum_{i=1}^{p} \frac{(b_{ei} + b'_{ei})}{2} + \sum_{j=1}^{p^2} v_j = R_e \qquad (4)$$

when:

$$\frac{\partial \pi}{\partial r_e} = 0$$

Since $\sum_{1}^{p} b_{ei} = 0$ and $\sum_{1}^{p^2} v_1 = 0$, we have:

$$p^2m + p^2r_e = R_e \quad \text{or} \quad r_e = \frac{R_e}{p^2} - \frac{T..}{2qp^2}$$

Thus the effect of each replication may be determined in terms of known quantities.

The effect of the differences between blocks within a group may be determined by differentiating $\pi$ with respect to the term $\frac{(b_{ei} - b'_{ei})}{2}$ and setting the result equal to zero.

Hence:

$$pr_e - pr'_e + p\frac{(b_{ei} - b'_{ei})}{2} + p\frac{(b_{ei} + b'_{ei})}{2} = B_{ei} - B'_{ei}$$

when:

$$\frac{\partial \pi}{\partial \left(\frac{b_i - b'_i}{2}\right)} = 0$$

in the general case, so that:

$$\frac{(b_{ei} - b'_{ei})}{2} = \frac{(B_{ei} - B'_{ei})}{2p} - \frac{(r_e - r'_e)}{2} = \frac{(B_{ei} - B'_{ei})}{2p} - \frac{(R_e - R'_e)}{2p^2} \qquad (5)$$

This relation is an integral part of component "a", due to block variance in the analysis of variance table, as will be shown, and is readily calculated from data.

At this point in the analysis, all information available has been obtained from the separate replications in each group, so they are now added together to form the group totals as indicated in Figure I. The blocks are now confounded and the equation for a plot yield in both groups may be given by:

$$Y_{hij} = \mu + \gamma_h + \beta_i + V_j + \epsilon_{hij} \qquad (6)$$

where the block and varietal effects have q times their original values and $\gamma_h$ will represent the group effect. This equation may be minimized in the same manner as (1). The double prime (") will be used to distinguish the estimates of these new parameters from similar ones made previously. To find the group effect, (6) may be rearranged and partially differentiated with respect to $\gamma_h$:

$$p^2m'' + p^2g''_h + p(b''_1 + b''_2 + \ldots + b''_p) + (v''_1 + v''_2 + \ldots + v''_{p2}) = G_h \qquad (7)$$

when:

$$\frac{\partial \pi''}{\partial g''_h} = 0$$

It should be noted that in the problem under discussion $G_h$ may be equal to X.. or Y..

Since the sum of the block effects is zero for each group, we have the following value for the group effect:

$$g_h = \frac{(G_h - p^2 m'')}{p^2} = \frac{(2G_h - T..)}{2p^2} \qquad (8)$$

when there are two groups.

The effects due to blocks may be obtained by partial differentiation of (6) rearranged and set equal to zero. We have, for any blokk in the first group:

$$pm'' + pg_h + pb''_i + (v''_{ip-p+1} + v''_{ip-p+2} + \ldots + v''_{ip}) = B''_i = X_{i.} \qquad (9)$$

and similarly for any block in the second group:

$$pm'' + pg_h + pb''_i + (v''_i + v''_{i+p} + \ldots + v''_{i+p(p-1)}) = B''_i = Y_{.i} \qquad (10)$$

when:

$$\frac{\pi''}{b''_i} = 0$$

Before solving (9) and (10) for the block effects it is necessary to find the varietal effect; which mayg be done by differentiating (6) rearranged with respect to $v''_j$. Thus the varietal effect for the first vasiety may be expressed:

$$2m'' + g_1 + g_2 + b''_1 + b''_{p+1} + 2v''_1 = V_1 \qquad (11)$$

when:

$$\frac{\pi''}{v_1} = 0$$

If the equations like (11) are summed through the pth variety, that is for one group, we have:

$$2pm'' + pb''_1 + (b''_{p+1} + b''_{p+2} + \ldots + b''_{2p}) + 2(v_1 + v_2 + \ldots + v_p) =$$

$$V_1 + V_2 + \ldots + V_p = X_{1.} + Y_{.1} \qquad (12)$$

Solving between (12) and (9):

$$b''_i = \frac{(X_{1.} - Y_{.1} - 2pg_1)}{p} = \frac{(pX_{1.} - pY_{.1} - 2G_1 + T..)}{p^2} \qquad (13)$$

The value thus obtained is free from group and varietal effects and is zero when summed for a group. The varietal effects may be solved now from (11) by substituting in the proper values of $b_i$. Then for any variety in the first block:

$$v''_j = 1/2\ v_j - \frac{p(X_{1.} - Y_{.i}) - 2G_1}{p^2} - \frac{p(Y_{i.} - X_{.1}) - 2G_2}{p^2} - \frac{3T..}{p^2} \qquad (14)$$

and the value of $v_j$ for any individual plot may be obtained from (14) by dividing by q. It may be noted that this is not true in the case of blocks for then the replications within a group are confounded. The design has now been solved for the effects of all the various components.

The analysis of variance may now be made using the relations just obtained for the various effects. It is often desired to test the significance of certain effects while simultaneously condeding other effects. In view of the preceeding arguments the following identity may be noted:

$$S\,Y_{eij}^2 = (T..)m + \sum_{e=1}^{2q} r_e R_e + \sum_{e=1}^{q} \sum_{i=1}^{p} \frac{(b_{ei} - b'_{ei})}{2} (B_{ei} - B'_{ei})$$

$$+ \sum_{e=1}^{q} \sum_{i=1}^{p} \frac{(b_{ei} + b'_{ei})}{2} (B_{ei} + B'_{ei}) + \sum_{j=1}^{p^2} v_j\ V_j + S(Y_{eij} - mz_1 - r_e z_{2e} -$$

$$\frac{(b_{ei} - b'_{ei})}{2} z_{3i} - \frac{(b_{ei} + b'_{ei})}{2} z_{4i} - vz_{5j})^2 \qquad (15)$$

Where each of the summation components in the first part of the equation represents a reduction in total variance due to the particular factor concerned. Therefore, each is really a partial regression coefficient measuring the specific importance of a variable.

The sum of squares attributed to regression in then:

$$(T..)m + \sum_{e=1}^{2q} r_e R_e + \sum_{e=1}^{q}\sum_{i=1}^{p} \left(\frac{b_{ei} - b'_{ei}}{2}\right)(B_{ei} - B'_{ei}) + \sum_{e=1}^{q}\sum_{i=1}^{p} \left(\frac{b_{ei} + b'_{ei}}{2}\right)$$

$$(B_{ei} + B'_{ei}) + \sum_{j=1}^{p^2} v_j V_j \qquad (16)$$

while the residual sum of squares or error is:

$$S(Y_{eij} - mz_1 - r_e z_{2e} - \left(\frac{b_{ei} - b'_{ei}}{2}\right) z_{3i} - \left(\frac{b_{ei} + b'_{ei}}{2}\right) z_{4i} - vz_{5j})^2 \qquad (17)$$

It is noted that (16) and (17) together make up the total variance of the population examined so that the latter may be obtained by subtraction.

From (4) and (16) the best estimate may be made of the expected value of the sum of squares in the regression due to replications. Noting that $m = \frac{T..}{2qp^2}$, $\sum_{e=1}^{2q} R_e = T..$, we have: $(T..)m = \frac{T..^2}{2qp^2}$ and:

$$\Sigma R_e r_e = \frac{1}{p^2} \sum_{e=1}^{2q} R_e^2 - \frac{(T..)^2}{2qp^2} \qquad (18)$$

which may be used in a table for the analysis of variance. The degrees of freedom attached to the replication sum of squares will then be (2q - 1)

From (5) and (16) the best estimate of the expected value of the sum of squares attributed to the differences in results of paired blocks maybe obtained. The expression for a pair of like replicates is found to be:

$$\frac{1}{2q}\sum_{e=1}^{q}\sum_{i=1}^{p} \left(\frac{B_{ei} - B'_{ei}}{2p}\right)^2 - \sum_{e=1}^{q} \left(\frac{R_e - R'_e}{2p^2}\right)^2 \qquad (19)$$

Using the notation that the difference between two block totals of two like replications in Group X be designated $B_{ix} - B'_{ix} = D_{ix}$ and similarly for Group Y, the total.

Expression for two groups to (19) may be written:

$$\frac{p\sum_{1}^{q}\sum_{1}^{p}D_x^2 + p\sum_{1}^{q}\sum_{1}^{p}D_y^2 - (\sum_{1}^{q}\sum_{1}^{p}D_x)^2 - (\sum_{1}^{q}\sum_{1}^{p}D_y)^2}{2p^2} \qquad (20)$$

This is the expression for:

$$\sum_{e=1}^{q}\sum_{i=1}^{p}\frac{(b_{ei} - b'_{ei})}{2}(B_{ei} - B'_{ei})$$

in equation (16), and is called component "a" of the sumof squares. It can be shown to have $2(p-1)$ degrees of freedom.

The next step is to obtain the best estimate of component "b" which corr-espondes to:

$$\sum_{e=1}^{q}\sum_{i=1}^{p}\frac{(b_{ei} + b'_{ei})}{2}(B_{ei} + B'_{ei})$$

term in (16). The general equation for the block effect is given by (9), which is necessarily confounded with variety effects due to the occurance of non-orthogonality in analysis. Therefore, the block totals must be lumped together for a group as indicated in Figure I. We may now, however, correct the block effect for varietal effect and replications effect by using the equation (13). This amounts to subtracting the varieties occurring in blocks for one group from the same varieties as they occur in a column for the other group, and finding the variances of these differences.

Adopting Cochran's[6] notation, we let the difference between a column total in Group X and the similar row total in Group Y by $X_{.i} - Y_{i.} = rhc_{yi}$. In the same manner, $Y_{.i} - X_{i.} = rhc_{xi}$. From (13) we might write the block effect in these terms and noting that $G_1 + G_2 = T..$, we obtain directly:

$$b''_{xi} = \frac{-prhc_x + \sum_{1}^{p} rhc_x}{p^2} \quad ; \quad b''_{yi} = \frac{-prhc_y + \sum_{1}^{p} rhc_y}{p^2} \qquad (21)$$

which is the effect in the respective groups. Since the like replications have been added, the block totals as given for a single replication before, are now

combined to form the group block total. The block effect term of the regression per block would then become:

$$\frac{1}{2q}\ \Sigma b_{xi}\ B_{xi} \quad \text{for Group X;} \qquad \frac{1}{2q}\ \Sigma b_{yi}\ B_{yi} \quad \text{for Group Y.}$$

For completeness, the block totals $B_{xi}$ and $B_{yi}$ are corrected for varieties by subtracting from the column total in the other group so that $B_{xi}$ becomes $-rhc_x$ and $B_{yi}$ becomes $-rhc_y$.

Hence, substituting the values of b and B, we have:

$$\frac{1}{2q}\ \Sigma\, b_{xi}\, B_{xi} = \frac{1}{2q}\ \frac{p\sum_1^p rhc_x^2 - (\Sigma\, rhc_x)^2}{p^2}; \quad \frac{1}{2q}\ \Sigma\, b_{yi}\, B_{yi} = \frac{1}{2q}\ \frac{p\sum_1^p rhc_y^2 - (\sum_1^p rhc_y)^2}{p^2}$$

The above equations may be combined now to give the total sum of squares attributable to blocks confounding like replications and correcting for varieties as:

$$\frac{p\sum_1^p rkc_x^2 + p\sum_1^p rkc_y^2 - (\sum_1^p rkc_x)^2 - (\sum_1^p rkc_y)^2}{2qp^2} =$$

$$\frac{p\,\Sigma\, rkc_x^2 + p\,\Sigma\, rkc_y^2 - 2(\sum_1^p rkc_x)^2}{2qp^2} \qquad (22)$$

This result agrees with that obtained by Nair[8] and is termed component "b" for the block regression. Like component "a" it is associated with 2(p- 1) degrees of freedom.

The varieties are of necessity confounded with replications and groups as well, so only variety totals for the entire experiment are used. It may now be noted that (14) could be written in terms of our new notation as:

$$v_j = \frac{1}{2p^2}\ (p^2 V_j + prkc_x + prkc_y - T..) \qquad (23)$$

Multiplying (23) by the variety total and summing, we obtain the regression sum of squares due to varieties alone in the form:

$$\frac{p^2 \sum_{1}^{p^2} V_j^2 - T..^2}{2qp^2} \qquad (24)$$

There are $(p^2 - 1)$ degrees of freedom allotted to varieties in an analysis.

The error term is found by subtracting the sums of squares due to the various known effects from the total variance of all the values, which is from the usual definition of variance:

$$\frac{2qp^2 \sum_{1}^{2qp^2} V_j^2 - T..^2}{2qp^2} \qquad (25)$$

with $2qp^2 - 1$ degrees of freedom.

In summary, the complete table for analysis of variance is given in general form in Figure III.

ANALYSIS OF VARIANCE

| Source of variation | Degrees of Freedom | Sum of Squares |
|---|---|---|
| Varieties | $p^2 - 1$ | $\frac{p^2 \sum_{1}^{p^2} V_j^2 - T..^2}{2qp^2}$ |
| Blocks: | | |
| Component "a" | $2(p - 1)$ | $\frac{p \sum_{1}^{p} D_{xi}^2 + p \sum_{1}^{p} D_{ji}^2 - (\sum_{1}^{p} D_{xi})^2 - (\sum_{1}^{p} D_{yi})}{2p^2 p}$ |
| Component "b" | $2(p - 1)$ | $\frac{p \sum_{1}^{p} rkc_{xi}^2 + p \sum_{1}^{p} rkc_{yi}^2 - 2(\sum rkc_{xi})^2}{2qp^2}$ |
| Replications | $2q - 1$ | $\frac{2q \sum_{I}^{2q} R_e^2 - T..^2}{2qp^2}$ |
| Error | $2qp^2 - p^2 - 4p - 2q + 5$ | by subtraction |
| Total | $2qp^2 - 1$ | $\frac{2qp^2 \sum_{1}^{2qp^2} v_j^2 - T..^2}{2qp^2}$ |

It has been prove d by Yates[8] that the significance of the sum of squares for any variate may be estimated by dividing it by the remaining sum of squares involving all variates except the one being tested. The ratio may then be tested against the appropriate value in a Snedecor's "F" table.

Although the design thus far has been solved for all variables, all the available information has not been utilized. It may be noted by referring to Figure I, that since block 1 of Group X contains the varieties 1 through p, and block 2 contains varieties p + 1 through 2p, the difference between the means of these two blocks is also an estimate of the difference between these two sets of varieties. The information obtained from such an inter-block consideration may be used provided the estimates derived from the comparison are proprely combined with the previous estimate. Since there are p plots contained in each block total, it is natural to expect a higher between block variance than within , for we have a situation analogous to the split plot design where the main block comparison usually exhibits a higher variance than do sub-plots. It is, therefore, necessary to estimate the error variance of the inter- and intra- block comparisons and to combine them to the best advantage. Both Yates and Cochran have found solutions by regarding the design as factorial, but a method will be developed here which will be along the lines of the analysis just discussed, and yet lead to the same general results.

The error for any plot may be regarded as the sum of two independent parts first a part e which varies from plot to plot with variance, say $\sigma_1^2$, and a part b which is constant for all plots in a given block. The latter error varies from block to block with a variance $\sigma_b^2$, and the error of the difference between two plots is $(e_{1i} - e_{1j})$. The error of a block total may then be written:

$$E_i = e_1 + e_2 + \ldots + e_p + pb_i$$

which has a variance of:

$$p\sigma_1^2 + p^2\sigma_b^2$$

The variance of the sum or difference of block totals for the groups is:

$$q\,2p\sigma_1^2 + q2p^2\,\sigma_b^2$$

To obtain component "a" the squares of the differences between the block totals are divided by 2qp to obtain the error per block. Thus, the mean square for component "a" is really an estimate of:

$$\sigma_1^2 + p\,\sigma_b^2 \qquad (26)$$

Component "b", however, is corrected for varieties by subtracting the coinciding variety totals of the other groups and not associated with a block. Since each e is independent of all other e's, as well as all the b's, the coefficient of $\sigma_1^2$ for the mean plot will be unity. The problem then resolves to finding the expected value of the inter-block variance. The expression for the regression sum of squares due to the component "b" block effect has been derived (22), and for our present prupose it will be permissable to consider only one group, multipy by two, and divide by the degrees of freedom to obtain the mean square for the error. The espression then becomes:

$$\frac{p\sum_1^p rhc_x^2 - (\sum_1^p rhc_x)^2}{2qp^2\,(p-1)} \qquad (27)$$

The error for $rhc_x$ may be written immediately as:

$$q\left[pb_1 - (b_{p+1} + b_{p+2} + \ldots + b_{2p})\right] \qquad (28)$$

since each group contains q replications. Squaring (28) from the aspect of the error only, we find the variance for $rhc_x^2$ is $qp\,(1+p)\,\sigma_b^2$. Summing over allvalues in a group and multiplying by p, the error of :

$$p\sum_1^p rhc_x^2$$

is:

$$qp^3\,(p+1)\,\sigma_b^2 \qquad (29)$$

The expected value of $\sum_{1}^{p} rhc_x$ of inter block error may be written now in terms of error as:

$$qp(b_1 + b_2 + \ldots + b_p) - qp\,(b_{p+1} + b_{p+2} + \ldots + b_{2p}).$$

and when this error is squared to become the variance of:

$$(\sum_{1}^{p} rhc_x)^2$$

we have:

$$q(p^3 + p^3)\,\sigma_b^2 = 2qp^3\,\sigma_b^2 \qquad (30)$$

When (29) and (30) are substituted in the numerator of (27), it is expressed in terms of the interblock variance. Adding the intra-block variance and reducing the expression, we establish that component "b" is an estimate of:

$$\sigma_i^2 + 1/2\ p\sigma_b^2 \qquad (31)$$

To summarize we now have the expected values for the following mean squares as shown in Figure III.

An estimate of $\sigma_i^2$ is available from the error and component "a" gives a direct estimate of the error mean squares between blocks since the difference between totals of pairs of blocks, contains the same set of varieties. This latter estimate can be improved by adjusting component "b" also to give an estimate of $\sigma_i^2 + p\,\sigma_b^2$.

For combining component "a" with component "b", it has been shown by Yates[5] and Cochran[6] that the loss in accuracy due to a simple average is very slight, and may be neglected. Hence equal weights may be used.

By averaging (26) and (31) it may be seen that the average mean square, B, of components "a" and "b" is an estimate of:

$$\sigma_i^2 + 3/4\ p\ \sigma_b^2$$

FIGURE III

ANALYSIS OF VARIANCE

| Source of Variation | D/F | Sum of Squares | Mean Square | Expected Mean Square |
|---|---|---|---|---|
| Blocks: | | | | |
| Component "a" | $2(p-1)$ | $\frac{\sum_{1}^{p} D_{xi}^2 + p\sum_{1}^{r} D_{j1}^2 - (\sum_{1}^{p} D_{xi})^2 - (\sum_{1}^{p} D_{yi})^2}{2p^2p}$ | $\frac{\sum_{1}^{p} D_{xi}^2 + p\sum_{1}^{r} D_{j1}^2 - (\sum_{1}^{p} D_{xi})^2 - (\sum_{1}^{p} D_{yi})^2}{2p^2p \quad 2(p-1)}$ | $\sigma_1^2 + p\,\sigma_b^2$ |
| Component "b" | $2(p-1)$ | $\frac{\sum_{1}^{p} t\mu c_{xi}^2 + p\sum_{1}^{p} r\mu c_{yi}^2 - 2(\sum r\mu c_{xi})^2}{2qp^2}$ | $\frac{\sum_{1}^{p} r\mu c_{xi}^2 + p\sum_{1}^{p} r\mu c_{yi}^2 - 2(\sum r\mu c_{xi})^2}{2qp^2 \quad 2(p-1)}$ | $\sigma_1^2 + 1/2\,p\,\sigma_b^2$ |
| Error | $2qp^2-p^2-4p-2q+5$ | by subtraction | by subtraction | $\sigma_1^2$ |

Since the intra-block error mean square, E, is an estimate of $\sigma_i^2$, an estimate of $\sigma_i^2 + 9\,\sigma_b^2$, the true variance between blocks is given by $E + 4(B - E) = \frac{(4B - E)}{3}$ and W' may be taken as $\frac{3}{(4B - E)}$.

Hence, the variance is now estimated from $4(P - 1)$ degrees of freedom instead of just $2(p - 1)$ degrees of freedom in case we had used only component "a", for estimating the inter-block variance. All varieties could now be corrected for inter-block as well as intra-block variance forwhich an elegant method has been developed by Harshbarger.[9]

## ESTIMATION OF THE MISSING PLOT

To estimate the value of a missing plot, the general procedure has been to minimize the error term and to use that value which gives the least error variance. This practice undoubtedly arose from the same sort of philosophy which inserts the mean for the missing value in a column of figures. The standard deviation is thus minimized andathe most probable value in the population isfound. Generalizing on the same scheme, Wishart and Allen minimized the error term to estimate missing values in designs of randomized blocks or Latin squares. Likewise, following the same course leads to minimizing the error term, E, in the analysis of variance forthe lattice design, which would then give a minimum value for the intra-block error.

Using "a" for the estimate of the missing value, Cornish[4] obtained the following equation in the terms of our notations:

$$a = \frac{2qrB''' + p^2V_i''' - p(X_{i.}''' + Y_{.i}''') - p(Y_{i.} - X_{.i}''') + (Y.. - X_{..}''')}{(p-1)(2qp-p-1)} \qquad (32)$$

where the triple prime (''') indicates totals of available data where the value, a, ismissing. The above equation holds for a value missing the the X Group, but a similar one may be written for the Y Group by simply reversing the X's and Y's. Equation (32) may be easily verified by assumming no variance and observing the following identities which may be summed and solved for a:

$$2qp(B''' + a) = T''' + a = 2qp^2a + a$$

$$p^2(V_i''' + a) = T''' + a$$

$$-p(X_{i.}''' + a + Y_{.i}) = -T''' - a$$

$$-p(Y_{i.} - X_{.i}''' - a) = -(Y.. - X''' - a) = 0$$

$$Y.. - X''' = a$$

It is found that the solution is (32), thus giving an estimate which will make

the error sum of squares or intra-block variance a minimum. We now come to the question of how to modify (32) to take into account the inter-block variance.

First consider what is expected of an estimate of a missing plot. Since the intra-block error hasalready been minimized, no amount of tampering with (32) will make it smaller. It seems reasonable to suppose, that if a value were missing from an experiment an estimate would be desired which would indicate as nearly as possible the missing value and use all the available information to make it comparable to the remaining data when all factors other than those to be considered and chance are excluded. With this aim in mind (32) may be adjusted to take into account the inter-block information to yield a closer, more accurate estimate of a missing value.

In the analysis of the design, it was seen that the variance due to blocks was divided into two parts designated components "a" and "b". Component "a" is the best direct estimate of inter-block variance, being a simple variation of block differences; while component "b" is really a term built on the deviation between a block total and its component varieties unconfounded with a block. To obtain the best estimate of the inter-block variance, then, itis clearly advantageous that component "b" be as small as possible. Therefore, to find a missing value in accordance with inter-block variance, let it be the value which makes component "bp" a minimum so that the mean square ascribed to the inter-block variance will then be as accurate as possible.

By denoting the missing value now by b and substituting in (22) we have:

$$\frac{p(\theta + \alpha^2 - 2\alpha b + b^2 + \lambda^2 - 2\lambda b + b^2) - (\tau^2 - 2\tau b + b^2 + \tau^2 - 2\tau b + b^2)}{2qp^2} \qquad (34)$$

Where: $\theta$ = sum of squares of all values of $rhc_x$ and $rhc_y$ not containing b

$$\alpha = Y_{.i} - X'''_{i.}$$

$$\lambda = Y_{i.} - X'''_{.i}$$

$$\tau = Y.. - X'''$$

When (34) is minimized by differentiating and setting equal to zero we obtain:

$$p\,(2\alpha + 2b - 2\lambda + 2b) - (-4\,\Sigma + 4b) = 0$$

$$b = \frac{p\,(\alpha + \lambda) - 2\Sigma}{2\,(p-1)} \tag{35}$$

or:

$$b = \frac{p(Y_{.i} + Y_{i.} - X''''_{i.} - X'''_{.i}) - 2\,(Y.. - X'''')}{2\,(p-1)}$$

which is the formulae for the missing plot in terms of inter-block variation only. The equation (35) may be verified in the same manner as used previously, by summing the following identities assuming no variance and solving for b.

$$\begin{aligned} p\,(Y_{.i} - X''''_{i.} - b) &= 0 \\ p\,(Y_{i.} - X'''_{.i} - b) &= 0 \\ -2\,(Y.. - X'''' - b) &= 0 \end{aligned} \tag{36}$$

To obtain an estimate of a missing plot using both the interblock and intra-block information, it is now necessary to combine (32) and (35) into a composite estimate using the proper weighting factors. Using the weighting factors for the two variances respectively, and letting the combined estimate be c, we have:

$$c = \frac{W}{W + W'} a + \frac{W'}{W + W'} b$$

which is at once both a manifest prodedure of interpolation and in accordance with Cochran's weighting procedure for obtaining estimates of main effects.

Before proceeding to the numerical example, let us first examine some of the consequences of using (37) in place of (32), to obtain the missing plot. In practice, (32) must be used alone to find an estimate of W and W'; after which (37) may be used if it is apparent that there is significant inter-block variance present. When the new value, c has been caluclated, it has been found in practice especially when the number of varieties tested is fairly large, that the effect on W and W' is absent or negligiable. If, however, the use of c in place of a does cause a significant change in W and W', further approximation may be carried out.

The extremes of (37) would occur in the cases where the error term is zero and where the sum of squares for blocks is zero. These conditions could happen simultaneously in which case the value would fit into a prefect law, and the question of block variance would cease to exist. When the error term is zero and block term is finite, we see that (37) reduces to (32) and c = a. It is impossible to have the third case where the error term is finite and ther blocks term is zero, because it has been seen from (26) and (31) that the error term is also involved in the block sum of square. However, compan-ent "a" or component "b" erroneously could be finite singly, andin these tirvial cases, the exact value of the missing plot could be obtained only by minimizing the component concerned. To cover such instances, it would be necessary to weight the estimate something like:

$$\frac{W}{W + W' + W''} a + \frac{W'}{W + W' + W''} b + \frac{W''}{W + W' + W''} c$$

where c would be the value obtained by minimizing component "a" with variance $\frac{1}{W''}$ . It has been found, however, that such versitility in venial instances is obtained at the expense of accuracy in the normal range of variances, because the intra-block error is often unduly increased by such estimates.

Therefore, use of (37) may be expected to give an estimate of a missing value by changing the error term as little as possible and at the same time decrease the effect of component "b" to a minimum. Thus the inter-block information is used as a modifier of the coarser estimate basedonly on the total experimental error.

---

## EXAMPLE: A CORN VARIETY EXPERIMENT

To illustrate the method of estimation just discussed, an experiment used by the Virginia Polytechnic Institute Experiment Station to test the yield of thirty-six varieties of corn in Lee County, Virginia, is presented. The yields varied greatly and consequently there was considerable inter-block variance. The experiment consisted of four replications, two randomized in Group X and two in Group Y. This gave, in all twenty-four blocks of six plots each arranged in the four replications of six blocks each.

Without indicating the random arrangements, Table I gives the actual individual plot yields for each variety in the four replications. The varieties are marked by the small number in the upper right corner of each plot. The block and replicate totals are also shown in the Table.

TABLE I

Plot Yields for Lattice Experiments of Lee County Corn

GROUP X

Replication I

| Blocks | | | | | | | Totals |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | |
| (1) | 23.4 | 20.1 | 25.5 | 25.8 | 40.5 | 12.3 | 147.6 |
| | 7 | 8 | 9 | 10 | 11 | 12 | |
| (2) | 31.2 | 36.6 | 18.6 | 21.3 | 13.2 | 27.3 | 148.2 |
| | 13 | 14 | 15 | 16 | 17 | 18 | |
| (3) | 14.7 | 29.4 | 32.7 | 1.5 | 20.1 | 19.8 | 118.2 |
| | 19 | 20 | 21 | 22 | 23 | 24 | |
| (4) | 23.1 | 18.3 | 11.7 | 78.3 | 12.0 | 18.9 | 102.3 |
| | 25 | 26 | 27 | 28 | 29 | 30 | |
| (5) | 33.0 | 24.6 | 31.8 | 28.2 | 15.6 | 17.7 | 150.9 |
| | 31 | 32 | 333 | 34 | 35 | 36 | |
| (6) | 27.0 | 21.0 | 30.6 | 49.2 | 37.8 | 34.8 | 200.9 |
| | | | | | | | 867.6 |

GROUP X (contd)

Replication II

| Blocks | | | | | | | Totals |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | |
| (1) | 59.7 | 66.6 | 30.9 | 41.4 | 43.5 | 19.2 | 261.3 |
| | 7 | 8 | 9 | 10 | 11 | 12 | |
| (2) | 126.3 | 130.8 | 126.3 | 66.0 | 96.6 | 63.9 | 609.9 |
| | 13 | 14 | 15 | 16 | 17 | 18 | |
| (3) | 122.1 | 106.8 | 126.9 | 111.9 | 104.1 | 94.8 | 666.6 |
| | 19 | 20 | 21 | 22 | 23 | 24 | |
| (4) | 26.4 | 21.3 | 26.1 | 15.3 | 21.9 | 16.8 | 127.8 |
| | 25 | 26 | 27 | 28 | 29 | 30 | |
| (5) | 42.9 | 30.6 | 51.6 | 30.9 | 27.3 | 26.3 | 219.6 |
| | 31 | 32 | 33 | 34 | 35 | 36 | |
| (6) | 45.0 | 17.1 | 40.2 | 58.8 | 42.3 | 40.8 | 244.2 |

2129.4

GROUP Y

Replication III

| Blocks | | | | | | | Totals |
|---|---|---|---|---|---|---|---|
| | 1 | 7 | 13 | 19 | 25 | 31 | |
| (1) | 42.6 | 75.0 | 27.3 | 66.3 | 57.6 | 47.4 | 316.2 |
| | 2 | 8 | 14 | 20 | 26 | 32 | |
| (2) | 32.1 | 22.5 | 36.6 | 22.5 | 21.3 | 12.0 | 138.0 |
| | 3 | 9 | 15 | 21 | 27 | 33 | |
| (3) | 44.1 | 31.2 | 25.8 | 59.4 | 47.1 | 52.5 | 260.1 |
| | 4 | 10 | 16 | 22 | 28 | 34 | |
| (4) | 32.1 | 41.4 | 11.4 | 49.5 | 30.3 | 51.0 | 215.7 |
| | 5 | 11 | 17 | 23 | 29 | 35 | |
| (5) | 29.7 | 9.6 | 34.5 | 44.7 | 12.6 | 26.1 | 157.2 |
| | 6 | 12 | 18 | 24 | 30 | 36 | |
| (6) | 28.2 | 44.4 | 52.5 | 42.6 | 48.9 | 43.5 | 260.1 |

1347.3

GROUP Y

Replication IV

| Blocks | | | | | | | Totals |
|---|---|---|---|---|---|---|---|
| | 1 | 7 | 13 | 19 | 25 | 31 | |
| (1) | 56.1 | 90.0 | 48.6 | 70.8 | 53.4 | 61.2 | 380.1 |
| | 2 | 8 | 14 | 20 | 26 | 32 | |
| (2) | 42.0 | 60.9 | 47.7 | 37.2 | 56.4 | 20.7 | 270.9 |
| | 3 | 9 | 15 | 21 | 27 | 33 | |
| (3) | 46.8 | 41.7 | 37.8 | 37.5 | 44.1 | 40.2 | 248.1 |
| | 4 | 10 | 16 | 22 | 28 | 34 | |
| (4) | 61.5 | 43.8 | 32.7 | 32.1 | 43.8 | 67.8 | 281.7 |
| | 5 | 11 | 17 | 23 | 29 | 35 | |
| (5) | 64.8 | 53.7 | 66.0 | 59.1 | 76.8 | 70.2 | 390.6 |
| | 6 | 12 | 18 | 24 | 30 | 36 | |
| (6) | 29.7 | 39.6 | 40.2 | 44.7 | 70.2 | 45.6 | 270.0 |
| | | | | | | | 1841.4 |

The like replications may now be combined to form the totals for each group. This is shown in Table II , which is the analog of Figure I.

TABLE II

Combinations of like Replications

GROUP X

| Blocks | | | | | | | Totals |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | |
| (1) | 83.1 | 86.7 | 56.4 | 67.2 | 84.0 | 31.5 | 408.9 |
| | 7 | 8 | 9 | 10 | 11 | 12 | |
| (2) | 157.5 | 167.4 | 144.9 | 87.3 | 109.8 | 91.2 | 758.1 |
| | 13 | 14 | 15 | 16 | 17 | 18 | |
| (3) | 136.8 | 136.2 | 159.6 | 113.4 | 124.2 | 114.6 | 784.8 |
| | 19 | 20 | 21 | 22 | 23 | 24 | |
| (4) | 49.5 | 39.6 | 37.8 | 33.6 | 33.9 | 35.7 | 230.1 |
| | 25 | 26 | 27 | 28 | 29 | 30 | |
| (5) | 75.9 | 55.2 | 83.4 | 59.1 | 42.9 | 54.0 | 337.5 |
| | 31 | 32 | 33 | 34 | 35 | 36 | |
| (6) | 72.0 | 38.1 | 70.8 | 108.0 | 80.1 | 75.6 | 444.6 |
| Totals | 574.8 | 523.2 | 552.9 | 468.6 | 474.9 | 402.6 | 2997.0 |

GROUP Y

| Blocks | | | | | | | Totals |
|---|---|---|---|---|---|---|---|
| | 1 | 7 | 13 | 19 | 25 | 31 | |
| (1) | 98.7 | 165.0 | 75.9 | 137.1 | 111.0 | 108.6 | 696.3 |
| | 2 | 8 | 14 | 20 | 26 | 32 | |
| (2) | 71.1 | 83.4 | 84.3 | 59.7 | 77.7 | 32.7 | 408.9 |
| | 3 | 9 | 15 | 21 | 27 | 33 | |
| (3) | 90.9 | 72.9 | 63.6 | 96.9 | 91.2 | 92.7 | 508.2 |
| | 4 | 10 | 16 | 22 | 28 | 34 | |
| (4) | 93.6 | 85.2 | 44.1 | 81.6 | 74.1 | 118.8 | 497.4 |
| | 5 | 11 | 17 | 23 | 29 | 35 | |
| (5) | 94.5 | 63.3 | 100.5 | 103.8 | 89.4 | 96.3 | 547.8 |
| | 6 | 12 | 18 | 24 | 30 | 36 | |
| (6) | 57.9 | 84.0 | 92.7 | 87.3 | 119.1 | 89.1 | 530.1 |
| Totals | 506.7 | 553.8 | 461.1 | 565.4 | 563.5 | 538.2 | 3188.7 |

Finally, the variety totals may be obtained by combining like varieties in Group X and Y to give us Table III.

TABLE III

VARIETY TOTAL YIELDS:

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| 181.8 | 157.8 | 147.3 | 160.8 | 178.5 | 89.4 |
| 7 | 8 | 9 | 10 | 11 | 12 |
| 322.5 | 250.8 | 217.8 | 172.5 | 173.1 | 175.2 |
| 13 | 14 | 15 | 16 | 17 | 18 |
| 212.7 | 220.5 | 223.8 | 157.5 | 224.7 | 207.3 |
| 19 | 20 | 21 | 22 | 23 | 24 |
| 186.6 | 99.3 | 134.7 | 115.2 | 137.7 | 123.0 |
| 25 | 26 | 27 | 28 | 29 | 30 |
| 186.9 | 132.9 | 174.6 | 133.2 | 132.3 | 173.1 |
| 31 | 32 | 33 | 34 | 35 | 36 |
| 180.6 | 70.8 | 163.5 | 226.8 | 176.4 | 164.7 |

The complete analysis of variance may now be made according to the equations set forth in Figure II. For the results shown in Table I, we have the following table for the analysis of variance shown as Table IV.

TABLE IV

Analysis of Variance--Lee County Corn

| Source of Variation | D/F | Sum of Squares | Mean Square |
|---|---|---|---|
| Replications: | 3 | 25,759.01 | 8,586.3366 |
| Component "a" | 10 | 25,741.59 | 2,574.159 |
| Component "b" | 10 | 14,782.62 | 1,478.262 |
| Blocks(eliminating var) | 20 | 40,524.21 | 2,026.2105 |
| Varieties(ignoring blocks) | 35 | 20,431.77 | 583.7648 |
| Error(intra-block) | 85 | 9,434.35 | 110.9924 |
| Total | 143 | 96,149.34 | |

From which we find W = 0.0000 and W' = 0.0004

If now we assume one plot to be missing, we may estimate its value by use of the formulae of the last section. As an extreme case, let us assume the yield for variety 16 to be missing in Replication 1 of Group X. First estimate the value from the intra-block information by using (32); then, we find the necessary quantities to be:

$q = 2$ $\quad p = 6$ $\quad B''' = 116.7$

$V'''_i = 156.0$ $\quad X'''_{i.} = 783.3$ $\quad Y_{i.} = 497.4$ $\quad Y.. = 3108.7$

$X'''_{.1} = 467.1$ $\quad Y_{.1} = 461.1$ $\quad X''' = 2995.5$

Whence:

$$a = \frac{2300.8 + 5616.0 - 7684.2 + 36.0 + 193.2}{5.7} = 11.3$$

Using this value in place of the one acutally obtained, we have the following table for the analysis of variance:

TABLE V

Analysis of Variance--Using 11.3 for Yield of Plot 16, Replicate 1 Group X

| Source of Variation | D/F | Sum of Squares | Mean Square |
|---|---|---|---|
| Replications | 3 | 25,391.43 | |
| Component "a" 10 | | 25,196.02 | |
| Component "b" 10 | | 15,082.31 | |
| Blocks(eliminating var) | 20 | 40,278.33 | 2,013.92 |
| Varieties(ignoring blocks) | 35 | 20,384.92 | |
| Error(intra-block) | 85 | 9,377.48 | 109.85 |
| Total | 144 | 95,432.16 | |

From Table V we find $W = .0091$ and $W' = .0004$

If, now, we solve for the missing value, using only inter-block information, we have according to (35):

$$b = \frac{-1751.4 - 386.4}{10} = -213.8$$

Combining this value with the estimate, a, and substituting in (37), we have:

$$c = 0.9574 \cdot 11.3 - 0.0421 \cdot 213.8 = 10.8 - 9.0 = 1.8$$

When c is substituted in for the missing value and an analysis of variance is made, we have Table VI:

TABLE VI

Analysis of Variance--Using 1.8 for Yield of Plot 16, Replicate 1 Group X

| Source of Variation | D/F | | Sum of Squares | Mean Squares | |
|---|---|---|---|---|---|
| Replications | 3 | | 25,747.70 | | |
| Component "a" | 10 | 25,724.69 | | | |
| Component "b" | 10 | 14,791.60 | | | |
| Blocks(eliminating var) | 20 | | 40,516.29 | 2,025.81 | W' = 0.0090 |
| Varieties(ignoring blocks) | 35 | | 20,429.64 | | |
| Error(intra-block) | 85 | | 9,430.92 | 110.95 | W = 0.0004 |
| Total | 143 | | 96,124.55 | | |

By comparing the Tables IV and V and VI, it can be seen that c is a much closer approximation to the actual conditions of the experiment than a. Also, the changes in the mean square for blocks as well as their respective weighting factors we negligiable.

Again , suppose that the value for plot 25, Replicate 1, Group X is missing. Using (32) we obtain:

$q=2$ $\quad V'''_i = 153.9$ $\quad X'''_{.i} = 541.8$

$p = 6$ $\quad X'''_i = 337.5$ $\quad Y_{.i} = 562.5$

$B''' = 117.8$ $\quad Y_{i.} = 696.3$ $\quad X''' = 2,964.0$

$Y.. = 3,188.7$

Whence:

$$a = \frac{2,827.2 + 5,540.4 - 6,202.8 - 124.2 + 224.}{5.17} = 26.7$$

When this value is used in place of the 33.0 actually found, the following table for the analysis of variance is obtained:

TABLE VII

Analysis of Variance-- using 26.7 for Yield of Plot 25, Replicate 1 Group X

| Source of Variation | D/F | | Sum of Squares | Mean Square | |
|---|---|---|---|---|---|
| Replications | 3 | | 25,997.42 | | |
| Component "a" | 10 | 25,595.66 | | | |
| Component "b" | 10 | 14,916.42 | | | |
| Blocks(eliminating var) | 20 | | 40,512.08 | 2,025.60 | W' = 0.000 |
| Varieties(ignoring blocks) | 35 | | 20,393.93 | | |
| Error(intra-block) | 85 | | 9,410.77 | 110.71 | W = 0.0090 |
| Total | 144 | | 96,314.20 | | |

If now we solvefor the missing plot using (35) we have:

$$b = \frac{2,277.0 - 449.4}{10} = 182.8$$

When the value of b is combined with a to find c, we have the following:

$$c = 0.9574 \cdot 26.7 + 0.0426 \cdot 182.8 = 33.3$$

which is much closer to the original value. When c is substituted back into the experimental results, we have the following analysis of variance:

TABLE VIII

Analysis of Variance--Using 33.3 for Yield of Plot 25, Replicate 1 Group X

| Source of Variation | D/F | | Sum of Squares | Mean Square | |
|---|---|---|---|---|---|
| Replications | 3 | | 25,747.70 | | |
| Component "a" | 10 | 25,748.67 | | | |
| Component "b" | 10 | 14,776.39 | | | |
| Blocks(eliminating var) | 20 | | 40,525.06 | 2,026.25 | W'= 0.000 |
| Varieties(ignoring blocks) | 35 | | 20,434.05 | | |
| Error(intra-block) | 85 | | 9,436.64 | 111.01 | W= 0.0090 |
| Total | 143 | | 96,143.45 | | |

By use of a closer approximation to the missing value, Table VIII is nearer to the actual figures in Table IV than is Table VII.

The modification based on inter-block information which is made on the existing intra-block formulae for estimating a missing block has thus shown itself capable of reproducing the experimental results more accurately than was previously possible.

Since the writing of this thesis, a very pertinent industrial problem has come to the attention of the author in which the lattice design has proved its efficiency. In a certain engineering process, a number of recordings are made automatically for each unit of production. The records are then measured by operators to determine various constants relating to the quality of the product. It was desired to find out the variability in the results caused by the difference between operators, and at the same time determine a method of correction for the difference. A lattice design has been set up which is identical in structure with the previous illustrative example whereby operations take place of blocks, production units take the

place of varieties, and replicate records on each unit are used for the replications.

Thus, the lattice design has found its way into industry, and may soon rightly take its place with the control charts among the tools of the industrial statistician.

## SUMMARY

The procedure introduced by Cornish[4] for estimating a missing value in a lattice design, has been extended here and modified to include not only intra-block information, but also inter-block information as well. The analysis of the lattice design has been presented with some simplification applicable to this type of lattice; thus, allowing a more uniform development than has appeared before.

The implications in the limit were examined for the use of the modification of intra-block estimation formula, and it was shown that the form used was an appropriate practical tool. An actual experiment was analysed, where certain plots were considered missing, and it was shown that the additional use of the inter-block information, as evidenced by the modified formula, would give a much better estimate of a missing value than had been possible previously in the presence of significant inter-block variation.

The lattice design has thus been made stable in the event of loss of any plot value. This design had been proven useful in industrial analysis, and its utility may now be protected by a more complete recovery of missing information.

## REFERENCES CITED


1. Yates, F. A New Method of Arranging Variety Trials Involving a Large number of Varieties. Journal Agricultural Science, 26: 424-455. 1936

2. Allen, F.E. and Wishart, J. A Method of Estimating the Yield of a Missing Plot in Field Experimental Work. Journal, Agircultural Science 20: 399-406. 1930

3. Yates, F. The Analysis of Replicated Experiments When Field Results Are Incomplete. Empire Journal, Experimental Agriculture, Vol. I. 129-142. 1933

4. Cornish, E. A. The Estimation of Missing Values in Quasi-Factorial Designs, Annals of Eugenics, 10:137-143. 1940

5. Yates, F. The Recovery of Inter-block Information in Variety Trials Arranged in Three Dimensional Lattices, Annals of Eugenics 9: 136-156. 1939

6. Cochran, W.G. The Analysis of Lattice and Triple Lattice Experiments in Corn Varietal Tests. Part II Mathematical Theory, Research Bulletin 281, Iowa State College of Agriculture, 1940.

7. Harshbarger, B. On the Analysis of a Certain Six by Six Four Group Lattice Design, Annals of Mathematical Statistics, 15: 307- 320. 1944

8. Yates, F. Orthogonal Functions and Tests of Significance in the Analysis of Variance. Supplemental Journal, Royal Statistical Society, 5: 177-180. 1938

9. Harshbarger, B. On the Analysis of a Certain Six by Six Four Group Lattice Design, Thesis in Library of the George Washington University, 1942.