MISSING PLOTS IN PARTIALLY BALANCED AND OTHER INCOMPLETE BLOCK DESIGNS

By M. N. Das

Indian Council of Agricultural Research, New Delhi

1. Introduction

Following the method of analysis of experiments with missing observations suggested by Yates (1933), Cornish (1940) deduced expressions for estimating missing values in the cases of one missing plot in balanced incomplete block design, Youden Square and quasifactorial designs. He did not give any expression for finding the standard errors of treatment differences.

An attempt has been made in this paper to give general expressions of (i) estimates of missing values, and (ii) variance of the difference between any two treatments, when two observations are missing in any manner whatsoever in the designs of balanced incomplete block, Youden Square, square lattice with m replications and partially balanced incomplete block. All the results in the case of one missing value can be obtained from those in the case of two missing plots. Though the method adopted is quite general to cover any number of missing plots, the results in the cases of more than two missing plots have not been attempted as these involve solution of equations with more than two unknowns. However, the method of writing the equations in the general case has been indicated.

2. Method of Analysis when Two Plots are Missing

In the analysis of incomplete block designs without recovery of inter-block information, the error sum of squares is obtained by subtracting the adjusted treatment sum of squares from the "within block" sum of squares. If there be two missing plots in any design and x and y are the substitutes for them, the expression containing x and y in 'within block' sum of squares is

$$(x^2 + y^2)\left(1 - \frac{1}{k}\right) - \frac{2xy\delta}{k} - \frac{2xB_x}{k} - \frac{2yB_y}{k},$$

where $\delta = 1$ or 0 according as both the missing plots are in the same block or not, B_x is the total of the block containing x taking 0 for the missing value, B_y similarly is the total of the block containing y and k is number of plots in a block.

Taking the adjusted total for the ith treatment $Q_i = T_i - V_i/k$ where T_i is the total of the observations having the ith treatment and V_i the sum of the totals of those blocks which contain the ith treatment and calling the plot where x has been substituted as the x plot and the other as the y plot, we can write

$$Q_i = T_i' + x\alpha_{xi} + y\alpha_{yi} - \frac{V' + x\beta_{xi} + y\beta_{yi}}{k}$$

where

$\alpha_{xi} = 1$ or 0 according as the ith treatment occurs in the x plot or not,

$\beta_{xi} = 1$ or 0 according as the ith treatment is in the block of the x plot or not,

$\alpha_{yi} = 1$ or 0 according as the ith treatment occurs in the y plot or not,

$\beta_{yi} = 1$ or 0 according as the ith treatment is in the block of the y plot or not.

T_i' and V_i' are the totals corresponding to T_i and V_i obtained by taking zero for the missing observations.

Q_i can also be expressed as

$$Q_i = Q_i' + x_i x + y_i y \text{ where } Q_i' = T_i' - \frac{V_i'}{k},$$

$$x_i = \left(\alpha_{xi} - \frac{\beta_{xi}}{k}\right) \text{ and } y_i = \left(\alpha_{yi} - \frac{\beta_{yi}}{k}\right).$$

As the function of Q's giving the adjusted treatment sum of squares is different for different designs, the expression containing x and y in the adjusted treatment sum of squares is also different for different designs and so for the different designs this has been obtained separately together with the estimates of x and y and other results.

Case I.—Balanced Incomplete block design.

The adjusted treatment sum of squares in this design

$$= \frac{\Sigma Q_i^2}{rE} \text{ where } E \text{ is the efficiency of the design.}$$

$$= \frac{\Sigma (Q_i' + x_i x + y_i y)^2}{rE}$$

$$= x^2 \frac{\Sigma x_i^2}{rE} + y^2 \frac{\Sigma y_i^2}{rE} + 2xy \frac{\Sigma x_i y_i}{rE} + 2x \frac{\Sigma x_i Q_i'}{rE}$$

$$+ 2y \frac{\Sigma y_i Q_i'}{rE} + \frac{\Sigma Q_i'^2}{rE}$$

Hence the terms involving x and y in the error sum of squares are

$$x^2 \left(\frac{k-1}{k} - \frac{\Sigma x_i^2}{rE}\right) + y^2 \left(\frac{k-1}{k} - \frac{\Sigma y_i^2}{rE}\right) - 2xy \left(\frac{\delta}{k} + \frac{\Sigma x_i y_i}{rE}\right)$$

$$- 2x \left(\frac{B_x}{k} + \frac{\Sigma x_i Q_i'}{rE}\right) - 2y \left(\frac{B_y}{k} + \frac{\Sigma y_i Q_i'}{rE}\right).$$

Solving the equations obtained by differentiating the above with respect to x and y, we get

$$x = \frac{\left(\frac{k-1}{k} - \frac{\Sigma y_i^2}{rE}\right)\left(\frac{B_x}{k} + \frac{\Sigma x_i Q_i'}{rE}\right) + \left(\frac{\delta}{k} + \frac{\Sigma x_i y_i}{rE}\right)\left(\frac{B_y}{k} + \frac{\Sigma y_i Q_i'}{rE}\right)}{\left(\frac{k-1}{k} - \frac{\Sigma x_i^2}{rE}\right)\left(\frac{k-1}{k} - \frac{\Sigma y_i^2}{rE}\right) - \left(\frac{\delta}{k} + \frac{\Sigma x_i y_i}{rE}\right)^2}.$$

$$y = \frac{\left(\frac{k-1}{k} - \frac{\Sigma x_i^2}{rE}\right)\left(\frac{B_y}{k} + \frac{\Sigma y_i Q_i'}{rE}\right) + \left(\frac{\delta}{k} + \frac{\Sigma x_i y_i}{rE}\right)\left(\frac{B_x}{k} + \frac{\Sigma x_i Q_i'}{rE}\right)}{\left(\frac{k-1}{k} - \frac{\Sigma x_i^2}{rE}\right)\left(\frac{k-1}{k} - \frac{\Sigma y_i^2}{rE}\right) - \left(\frac{\delta}{k} + \frac{\Sigma x_i y_i}{rE}\right)^2}.$$

As x_i can take only k non-zero values, *viz.*, $1 - 1/k$ for the affected treatment and $-1/k$ for each of the other $k-1$ treatments in the block of x,

$$\Sigma x_i^2 = \frac{k-1}{k}.$$

Similarly, $\Sigma y_i^2 = \frac{k-1}{k}$.

So by putting

$$a = \left(\frac{k-1}{k}\right)\left(1 - \frac{1}{rE}\right), \quad c = \frac{\delta}{k} + \frac{\Sigma x_i y_i}{rE}$$

and $\Delta = a^2 - c^2$, x and y can be written as

$$x = \frac{a\left(\frac{B_x}{k} + \frac{\Sigma x_i Q_i'}{rE}\right) + c\left(\frac{B_y}{k} + \frac{\Sigma y_i Q_i'}{rE}\right)}{\Delta}$$

$$y = \frac{a\left(\frac{B_y}{k} + \frac{\Sigma\, y_i Q_i'}{rE}\right) + c\left(\frac{B_x}{k} + \frac{\Sigma\, x_i Q_i'}{rE}\right)}{\Delta}$$

If the plots are missing in the same block, $B_x = B_y$, $\delta = 1$ and

$$\Sigma\, x_i y_i = \frac{-2\,(k-1)}{k^2} + \frac{k-2}{k^2} = -\frac{1}{k}.$$

If the plots are missing in two blocks having p treatments common, then $\delta = 0$ and $\Sigma x_i y_i = p/k^2 - q/k$ where $q = 0$, 1 or 2 according as none, one or two of the affected treatments are among the p common treatments, provided, however, that the same treatment is not missing in both the blocks. In the latter case, $\Sigma x_i y_i = 1 - 2/k + p/k^2$.

Analysis of the data thus completed will provide the unbiassed error sum of squares. The unbiassed adjusted treatment sum of squares can then be obtained by subtracting the error sum of squares from the 'within block' sum of squares obtained from the incomplete data. This method is the same for all designs excepting Youden Square once the missing values are estimated.

If t_i and t_j be any two treatments, it is well known that the variance of $(t_i - t_j)$ can be obtained by first expressing $(t_i - t_j)$ as $k_i T_i' - k_j T_j'$ + terms not containing T_i' and T_j' and then taking the variance as $(k_i + k_j)\,\sigma^2$. As $Q_i' = T_i' - V_i'/k$, the variance can also be obtained by finding the coefficients of Q_i' and Q_j' and then subtracting them.

As $t_i = Q_i/rE$ and can be obtained by substituting the values of x and y, we have

$$t_i - t_j = \frac{Q_i' - Q_j'}{rE} + \frac{x\,(x_i - x_j)}{rE} + \frac{y\,(y_i - y_j)}{rE}.$$

Now, by substituting for x and y and collecting the coefficients of Q_i' and Q_j', the variance comes out to be

$$\text{Var}\,(t_i - t_j) = \frac{\sigma^2}{rE}\left[2 + \frac{(k-1)\,(rE-1)}{k\Delta r^2 E^2}\{(x_i - x_j)^2 + (y_i - y_j)^2\} + \frac{2\,(\delta rE + k\Sigma x_i y_i)\,(x_i - x_j)\,(y_i - y_j)}{k\Delta r^2 E^2}\right].$$

As the values of $(x_i - x_j)$ and $(y_i - y_j)$ for particular values of i and j depend on how they occur in the affected blocks, it is not always possible to express the above expression of variance in terms of the parameters of the design.

If there be only one missing plot, say x, x_i's will remain the same but y_i's will be zero, so that

$$x = \frac{\frac{B_x}{k} + \frac{\Sigma x_i Q_i'}{rE}}{\frac{k-1}{k}\left(1 - \frac{1}{rE}\right)} = \frac{rE\, B_x + k\,\Sigma x_i Q_i'}{(k-1)(rE-1)}.$$

This can also be written as

$$\frac{rEB_x - \Sigma\, Q_l' + kQ_m'}{(k-1)(rE-1)},$$

where m stands for the affected treatment and l extends over all the treatments in the affected block including m. Due to change of notations this expression has come out in a simpler form than that obtained by Cornish (1940).

When $v = k$, E becomes unity and the estimate reduces to

$$\frac{rB_x + kT_i' - G}{(r-1)(k-1)},$$

the well-known expression in the case of the randomised block design. Variance $(t_i - t_j)$ becomes

$$\frac{\sigma^2}{rE}\left\{2 + \frac{k\,(x_i - x_j)^2}{(k-1)(rE-1)}\right\}.$$

This reduces to

(i) $\frac{\sigma^2}{rE}\left\{2 + \frac{k}{(k-1)(rE-1)}\right\}$ when i is affected and j in the affected block.

(ii) $\frac{\sigma^2}{rE}\left\{2 + \frac{k-1}{k\,(rE-1)}\right\}$ when i is affected and j not in the affected block.

(iii) $\frac{\sigma^2}{rE}\left\{2 + \frac{1}{k\,(k-1)(rE-1)}\right\}$ when i is in the affected block but not affected itself and j not in the affected block.

(iv) $\frac{2\sigma^2}{rE}$ when none or both of i and j are in the affected block.

As it should be the variance of the treatment difference has increased or remained the same as in the non-missing case and the amount of increase is easily obtainable from the expressions.

Case II.—Youden Square.

In order to account for the row sum of squares which is necessary in this design, another function of x and y, *viz.*,

$$\left(\frac{x^2+y^2}{b}\right)\left(1-\frac{1}{k}\right)+\frac{2xy}{b}\left(\delta'-\frac{1}{k}\right)+\frac{2x}{b}\left(R_x-\frac{G}{k}\right)+\frac{2y}{b}\left(R_y-\frac{G}{k}\right),$$

is to be subtracted from the expression containing x and y in the error sum of squares derived in the case of balanced incomplete block design. Here $\delta' = 1$ or 0 according as both the missing values are in the same row or not, R_x is the total of the row containing x, R_y the total of the row containing y and G is the grand total—all obtained by taking zero for the missing values.

It can now be shown easily that the different results in Youden Square can be obtained from those of B.I.B. design by writing

$$\left(1-\frac{1}{k}\right)\left(1-\frac{1}{rE}-\frac{1}{b}\right) \text{ for } \left(1-\frac{1}{k}\right)\left(1-\frac{1}{rE}\right),$$

$$\frac{\delta}{k}+\frac{\Sigma\, x_i y_i}{rE}+\frac{\delta'}{k}-\frac{1}{bk} \text{ for } \frac{\delta}{k}+\frac{\Sigma\, x_i y_i}{rE},$$

$$\frac{B_x}{k}+\frac{\Sigma\, x_i Q_i'}{rE}+\frac{R_x}{b}-\frac{G}{bk} \text{ for } \frac{B_x}{k}+\frac{\Sigma\, x_i Q_i'}{rE},$$

and

$$\frac{B_y}{k}+\frac{\Sigma\, y_i Q_i'}{rE}+\frac{R_y}{b}-\frac{G}{bk} \text{ for } \frac{B_y}{k}+\frac{\Sigma\, y_i Q_i'}{rE}.$$

In the case of one missing plot, x becomes

$$\frac{rE\,(bB_x + kR_x - G) - b\Sigma Q_i' + bkQ_i'}{(k-1)\,(brE - rE - b)},$$

which reduces to the corresponding expression in B.I.B. design by first dividing both the numerator and the denominator by b and then putting $1/b$ equal to zero.

If E_1 be the error S.S. from the completed data and E_2 that from the block X row table obtanied by taking the missing values into account, then $E_2 - E_1$ gives the unbiassed treatment S.S., and the error S.S. is provided by E_1.

$$\text{Var}\,(t_i - t_j) = \frac{\sigma^2}{rE}\left\{2 + \frac{bk\,(x_i - x_j)^2}{(k-1)\,(brE - rE - b)}\right\}$$

$$= \frac{\sigma^2}{rE}\left\{2 + \frac{k\,(x_i - x_j)^2}{(k-1)\,(k-2)}\right\},$$

since by virtue of $b = v$, $(brE - rE - b)$ becomes $b\,(k-2)$.

The values of $(x_i - x_j)$ in terms of the parametres of the design remains the same as in the case of B.I.B. design.

Case III.—Partially balanced incomplete block design.

In a partially balanced incomplete block design with two associate classes, the treatment sum of squares is obtained from ΣtQ, where

$$t_i = c_1 Q_i + c_2 \Sigma Q_{i1},$$

$$c_1 = \frac{kB_{22}}{\Delta}, \quad c_2 = -\frac{kB_{12}}{\Delta},$$

$$\Delta = A_{12}B_{22} - A_{22}B_{12},$$

$$A_{12} = r\,(k-1) + \lambda_2,$$

$$B_{12} = (\lambda_2 - \lambda_1),$$

$$A_{22} = (\lambda_2 - \lambda_1)\, p_{12}{}^2$$

$$B_{22} = r\,(k-1) + \lambda_2 + (\lambda_2 - \lambda_1)\,(p_{11}{}^1 - p_{11}{}^2).$$

ΣQ_{i1} is the sum of Q's corresponding to those treatments which are the first associates of the ith treatment. For convenience of analysis, the smaller of the two groups of treatments corresponding to the two associate classes will be taken to form the first associate class.

In presence of two missing plots,

$$t_i = c_1\,(Q_i' + x_i x + y_i y) + c_2\,(\Sigma Q_{i1}' + x\Sigma x_{i1} + y\Sigma y_{i1}),$$

where Σx_{i1} and Σy_{i1} are respectively the summations of x_i's and y_i's corresponding to those treatments which are the first associates of the ith treatment.

Putting $c_1 Q_i' + c_2 \Sigma Q_{i1}' = t_i'$, $c_1 x_i + c_2 \Sigma x_{i1} = X_i$ and $c_1 y_i + c_2 \Sigma y_{i1} = Y_i$, so that t_i' is the estimate of t_i obtained by taking zero for the missing values, we can write

$$t_i = t_i' + xX_i + yY_i.$$

Thus, the treatment sum of squares ΣtQ becomes

$$x^2 \Sigma x_i X_i + y^2 \Sigma y_i Y_i + xy\,(\Sigma x_i Y_i + \Sigma X_i y_i) + x\,(\Sigma x_i t_i' + \Sigma X_i Q_i')$$
$$+ y\,(\Sigma y_i t_i' + \Sigma Y_i Q_i') + \text{terms without } x \text{ and } y.$$

It follows easily from the properties of P.B.I.B. design that

$$\Sigma X_i y_i = \Sigma x_i Y_i,\ \Sigma x_i t_i' = \Sigma X_i Q_i' \text{ and } \Sigma y_i t_i' = \Sigma Y_i Q_i'.$$

Hence the terms involving x and y in the error sum of squares can be written as

$$x^2 \left\{\frac{k-1}{k} - \Sigma x_i X_i\right\} + y^2 \left\{\frac{k-1}{k} - \Sigma y_i Y_i\right\} - 2xy \left\{\frac{\partial}{k} + \Sigma x_i Y_i\right\}$$
$$- 2x \left\{\frac{B_x}{k} + \Sigma X_i Q_i'\right\} - 2y \left\{\frac{B_y}{k} + \Sigma Y_i Q_i'\right\}.$$

Putting

$$\frac{k-1}{k} - \Sigma x_i X_i = A,\ \frac{k-1}{k} - \Sigma y_i Y_i = B,\ \frac{\partial}{k} + \Sigma x_i Y_i = C$$

and $AB - C^2 = \Delta$, the estimate of x and y can thus be obtained from

$$x = \frac{B\left(\frac{B_x}{k} + \Sigma X_i Q_i'\right) + C\left(\frac{B_y}{k} + \Sigma Y_i Q_i'\right)}{\Delta}$$

$$y = \frac{A\left(\frac{B_y}{k} + \Sigma Y_i Q_i'\right) + C\left(\frac{B_x}{k} + \Sigma X_i Q_i'\right)}{\Delta}$$

If E denotes the error S.S. from the completed data and B the 'within block S.S.' from the incomplete data, then $(B - E)$ gives the unbiassed treatment S.S. and E, the error S.S. The treatment S.S. can be obtained directly also from $\Sigma Q_i''\,(t_i' + xX_i + yY_i)$ where Q_i'' is the adjusted total obtained from the incomplete data as they are and not taking the missing values as zero. The estimate of the difference between any two treatments can be obtained from

$$(t_i - t_j) = (t_i' - t_j') + x\,(X_i - X_j) + y\,(Y_i - Y_j).$$

Substituting for x and y and collecting coefficients of Q_i' and Q_j', the variance of the difference between the two treatments is obtained as

$$\text{Var}\,(t_i - t_j) = V + \frac{(X_i - X_j)}{\Delta}\{B\,(X_i - X_j) + C\,(Y_i - Y_j)\}$$

$$+ \frac{(Y_i - Y_j)}{\Delta}\{A\,(Y_i - Y_j) + C\,(X_i - X_j)\},$$

where V is the variance of $(t_i - t_j)$ in the non-missing case,

$$= V + \frac{1}{\Delta}\{B\,(X_i - X_j)^2 + A\,(Y_i - Y_j)^2$$

$$+ 2C\,(X_i - X_j)\,(Y_i - Y_j)\}.$$

The results in the case of one missing value, say x, can be obtained from the above by taking y_i to be zero.

Thus,

$$x = \frac{\dfrac{B_x}{k} + \Sigma Q_i' X_i}{\dfrac{k-1}{k} - \Sigma x_i X_i} = \frac{B_x + k\Sigma t_i' x_i}{(k-1) - k\Sigma x_i X_i}, \text{ as } \Sigma X_i Q_i' = \Sigma t_i' x_i.$$

As in this design also, x_i takes the value $1 - 1/k$ for the affected treatment, $-1/k$ for each of the other $k - 1$ treatments in the affected block and zero for all the other treatments, the estimate of x can also be written as

$$x = \frac{B_x - \Sigma t_l' + k t_m'}{(k-1) - k\Sigma x_i X_i},$$

where l and m have the same meaning as in the corresponding expressions of B.I.B. design and Youden Square.

$$\text{Variance}\,(t_i - t_j) = V + \frac{k\,(X_i - X_j)^2}{(k-1) - k\Sigma x_i X_i}.$$

As t does not seem possible to evaluate in general X_i in terms of the parameters of the design, the variance also cannot be expressed purely in terms of the parameters of the design.

Though all the above results for the partially balanced incomplete block design have been derived for designs with two associate classes only, the results obtained are true in general for designs with any number of associate classes excepting for the new meaning of X_i, Y_i and t_i', which, in the general case of s associates, are to be obtained from

$$X_i = C_1 x_i + C_2 \Sigma x_{i1} + \cdots + C_s \Sigma x_{i,s-1},$$

$$Y_i = C_1 y_i + C_2 \Sigma y_{i1} + \cdots + C_s \Sigma y_{i,s-1},$$

and

$$t_i' = C_1 Q_i' + C_2 \Sigma Q_{i1}' + \cdots + C_s \Sigma Q_{i,s-1}',$$

where C_1, C_2, $\cdots$ are the coefficients of Q_i, ΣQ_{i1}, etc., in the solution for t_i, Σx_{ij} is the sum of those x_i's which correspond to the treatments which are the jth associates of the ith treatment; Σy_{ij} also denotes similar summation over the y_i's.

Case IV.—Lattice Designs.

Square lattice designs with m replications is a particular type of partially balanced incomplete block design with two associate classes. In this case, it is possible to express X_i's and all other results depending on them in terms of the parameter of the design, namely,

$$C_1 = \frac{1}{m} + \frac{1}{k(m-1)}, \quad C_2 = \frac{1}{mk(m-1)}$$

$$\Sigma x_i^2 = \frac{k-1}{k}, \qquad \Sigma x_i \Sigma x_{i1} = -\frac{k-1}{k},$$

so that

$$\Sigma x_i X_i = C_1 \Sigma x_i^2 + C_2 \Sigma x_i \Sigma x_{i1}$$

$$= \frac{k-1}{k}(C_1 - C_2)$$

$$= \frac{k^2-1}{mk^2}.$$

X_i takes the value $(k^2-1)/k^2m$ for the affected treatment, $-(k+1)/k^2m$ for each of the other $k-1$ treatments in the affected block,

$$\frac{k-m+1}{mk^2(m-1)}$$

for each of the other treatments which occur with the affected treatment in other blocks and $-1/mk^2$ for all treatments which do not occur with the affected treatment in a block.

When two plots are missing, $\Sigma x_i Y_i$ cannot be expressed in terms of the parameters of the design unless the contents of the blocks in which the plots are missing are known. When one plot is missing, the various results have been given below in terms of the parameters of the design, as in B.I.B. design, namely,

$$x = \frac{km\,(B_\alpha - \Sigma t_i' + kt_m')}{(k-1)\,(mk-k-1)}$$

$$\text{Var}\,(t_i - t_j) = V + \frac{mk^2\sigma^2\,(X_i' - X_j)^2}{(k-1)\,(mk-k-1)}$$ where V is the variance of the difference in the non-missing case,

$$= V + \frac{(k+1)^2\,\sigma^2}{m\,(k-1)\,(mk-k-1)}$$ when i is the affected treatment and j occurs in the affected block,

$$= V + \frac{(km-k-1)\,\sigma^2}{m\,(k-1)\,(m-1)^2},$$ when i is affected and j occurs with i in a block other than the affected block,

$$= V + \frac{k^2\sigma^2}{m\,(k-1)\,(mk-k-1)},$$ when i is the affected treatment and j does not occur with i in any block.

$$= V + \frac{m\sigma^2}{(m-1)^2\,(k-1)\,(mk-k-1)},$$ when i is any treatment in the affected block and is not affected itself and j any other treatment in the other blocks with the affected treatment,

$$= V + \frac{\sigma^2}{m\,(k-1)\,(mk-k-1)},$$ when i is any treatment in the affected block, but not affected itself and j any treatment not occurring with the affected treatment in any block,

$$= V + \frac{\sigma^2}{m(m-1)^2(k-1)(mk-k-1)},$$ when i is any treatment occurring with the affected treatment in any block other than the affected one and j any other treatment not occurring with the affected treatment in any block,

$= V$, when both i and j occur together in a block where the affected treatment occurs, or they do not occur at all in any of these blocks.

3. General Case of any Number of Missing Plots

When more than two plots are missing, the various results cannot, in general, be expressed as simple algebraic expressions. The equations leading to the estimates of the missing value can be written in the general case also without much involvement. Such equations have been indicated below in the case of n missing plots substituted by m_i $(i = 1, 2 \cdots n)$ in the partially balanced incomplete block design. The equation in the case of other designs excepting Youden Square can be easily obtained from those of partially balanced incomplete block design. From the equations in balanced incomplete block design, those of Youden Square can be obtained by making the adjustments indicated in the case of two missing plots.

If

$\alpha_{m_{ij}} = 1$ or 0 according as the jth treatment occurs in the plot of m_i or not,

and

$\beta_{m_{ij}} = 1$ or 0 according as the jth treatment occurs in the block of m_i or not.

then, denoting

$$\left(\alpha_{m_{ij}} - \frac{\beta_{m_{ij}}}{k}\right) by\ m_{ij}$$

and

$$C_1 m_{ij} + C_2 \Sigma m_{ij1} + C_3 \Sigma m_{ij2} + \cdots, \text{ by } M_{ij},$$

where Σm_{ij}s is the summation of m_{ij}'s corresponding to those treatments which are the Sth associates of the jth treatment, the equations can be written as

$$m_i\left(\frac{k-1}{k}-\sum_j m_{ij}M_{ij}\right)-\sum_{l\neq i} m_l\left(\frac{\delta_{il}}{k}+\sum_j m_{lj}M_{ij}\right)$$
$$=\left(\frac{B_i}{k}+\sum_j M_{ij}Q_j'\right)\quad (i=1,2,\cdots n),$$

where $\delta_{il}=1$ or 0 according as both the m_i and m_l plots occur in the same block or not and B_i is the total of the block containing the m_i plot obtained by taking zero for the missing values.

In several particular cases, algebraic solutions of the equations are available. Thus, in a B.I.B. design, if the plots are missing each in a different block, such that there are no common treatments among the affected blocks, the equations become

$$\left(\frac{k-1}{k}\right)\left(1-\frac{1}{rE}\right)m_i=\frac{B_i}{k}+\frac{\sum_j m_{ij}Q_j'}{rE}\quad (i=1,2,\ldots n).$$

This shows that each of the missing values in such cases can be estimated independently of others. The variance of the difference between any two treatments comes out to be

$$\text{Var}\,(t_j-t_j')=\frac{\sigma^2}{rE}\left\{2+\frac{k\sum_i(m_{ij}-m_{ij}')^2}{(k-1)\,(rE-1)}\right\}.$$

If, again, all the missing values are in the same block provided the block is not completely missing, the equations reduce to

$$\frac{k-1}{k}\left(1-\frac{1}{rE}\right)m_i-\sum_{l\neq i} m_l\left(\frac{1}{k}-\frac{1}{krE}\right)$$
$$=\frac{B_i}{k}+\frac{\sum_j m_{ij}Q_j'}{rE}\quad (i=1,2,\cdots n).$$

Solving these equations,

$$m_i=\frac{rEB_i+(k-n)\,Q_i'-\Sigma Q_l'}{(k-n)\,(rE-1)},$$

where $\Sigma Q_l'$ denotes the summation over the Q's corresponding to those treatments in the affected block which are not affected themselves and Q_i' corresponds to the treatment of m_i.

The variance of the difference between any two treatments is given by

$$\text{Var}\,(t_j-t_{j'})=\frac{\sigma^2}{rE}\left[2+\frac{\sum_i(m_{ij}-m_{ij}')^2}{(rE-1)}+\frac{\left\{\sum_i(m_{ij}-m_{ij}')\right\}^2}{(k-n)\,(rE-1)}\right]$$

In lattice designs also, the equations are solvable when plots are missing, as in the above two cases. In partially balanced incomplete block designs with two associate classes such that one of the λ's is zero, the general equations admit of algebraic solution if all the missing plots are in the same block but the block is not completely missing. If a block be completely missing in such designs, the data can be analysed by a method given by Zelen (1954).

4. An Example

The method of analysis of partially balanced incomplete block design with two associate classes having two missing plots, has been illustrated below. The data analysed were obtained in the following manner.

Eight numbers, *viz.*, 4, 5, 7, 9, 10, 12, 13 and 15 were taken to represent in order the effect of the eight treatments $v_1, v_2, \cdots v_8$. Each of the numbers was repeated five times to give in all 40 numbers. These 40 numbers were then made into eight groups to form eight blocks of five each of the partially balanced incomplete block design $b = v = 8$, $r = k = 5$, $\lambda_1 = 4$, $\lambda_2 = 2$, $n_1 = 3$, $n_2 = 4$, $p^1_{11} = 2$. Next, eight numbers, *viz.*, 5, 2, 3, 8, 7, 1, 5 and 4 were taken and the first of them was added to each of the five numbers in the group forming the first block; the second number was similarly added to each of the five numbers in the second block and so on for other blocks, so that the 40 numbers thus obtained represented the block and treatment effects. Next, 40 numbers inbetween 3 and —4 were chosen at random so as to make their total zero and then added one to each of the 40 numbers representing the block and treatment effects. The numbers thus obtained formed the data and are shown below. The numbers of the treatments are given in brackets beside the data.

Blocks						Block totals	Treatment totals
1	$\frac{7}{x}$(1)	10(2)	15(3)	13(4)	14(5)	$52+x$	$37+x$
2	7(1)	8(2)	8(3)	9(4)	$\frac{14}{y}$(6)	$32+y$	41
3	10(1)	5(2)	12(3)	12(4)	17(7)	56	66
4	11(1)	9(2)	17(3)	20(4)	22(8)	79	66
5	9(1)	13(5)	20(6)	20(7)	21(8)	83	62
6	9(2)	9(5)	14(6)	15(7)	17(8)	64	$68+y$
7	14(3)	12(5)	18(6)	18(7)	21(8)	83	87
8	12(4)	14(5)	16(6)	17(7)	21(8)	80	102

The first plot in the first block and the last plot in the second block, which are marked by x and y respectively, have been taken to be missing. The expected values in these two plots are 9 for x and 14 for y.

For this design, $C_1 = 45/192$ and $C_2 = 5/192$.

The other expressions required to obtain the missing values, etc., can be obtained from the following table:—

Treatment No.	Adjusted total (Q_i')	x_i	y_i	Σx_{i1}	Σy_{i1}	ΣQ_{i1}	$t_i' \times 192$	$X_i \times 192$	$Y_i \times 192$
1	−23·4	·8	−·2	−·6	−·6	− 3·8	−1072	33	−12
2	−15·6	−·2	−·2	·4	−·6	−11·6	− 760	−7	−12
3	5·6	−·2	−·2	·4	−·6	−32·8	88	−7	−12
4	6·2	−·2	−·2	·4	−·6	−33·4	112	−7	−12
5	−10·4	−·2	0	0	·8	37·6	− 280	−9	4
6	− 0·4	0	·8	−·2	0	27·6	120	−1	36
7	13·8	0	0	−·2	·8	13·4	688	−1	4
8	24·2	0	0	−·2	·8	3·0	1104	−1	4

From the table,

$$\Sigma X_i Q_i' = \Sigma t_i' x_i = -689\cdot6/192$$

$$\Sigma Y_i Q_i' = \Sigma t_i' y_i = 422\cdot4/192$$

$$\Sigma x_i Y_i = \Sigma X_i y_i = -3\cdot2/192$$

$$\Sigma t_i' Q_i' = 77203\cdot2/192, \ \Sigma x_i X_i = 32\cdot4/192, \ \Sigma y_i Y_i = 38\cdot4/192$$

Checks are obtained from the totals of the different columns of the table, each of which should be zero. Also $\Sigma X_i Q_i'$, $\Sigma Y_i Q_i'$, $\Sigma x_i Y_i$ should respectively be equal to $\Sigma t_i' x_i$, $\Sigma t_i' y_i$, $\Sigma X_i y_i$.

Substituting the above in the expressions for x and y, we get $x = 10\cdot41$ and $y = 14\cdot04$.

The variance of the difference between any two treatments can be obtained easily with the help of the last two columns of the table. Actually there are 11 different types of variances as against two in the non-missing case. The adjusted treatment S.S. from the completed data is obtained from

$$\Sigma (t_i' + xX_i + yY_i)(Q_i' + xx_i + yy_i)$$

$$= \Sigma t_i' Q_i' + x^2 \Sigma x_i X_i + y^2 \Sigma y_i Y_i + 2x \Sigma x_i t_i' + 2y \Sigma y_i t_i' + 2xy \Sigma x_i Y_i$$

$$= 441\cdot94.$$

The within block S.S. from the completed data $= 515\cdot35$

Hence the unbiassed error S.S. $= 515\cdot35 - 441\cdot94 = 73\cdot41$

The within block S.S. from the completed data $= 480\cdot80$.

Hence the unbiassed treatment S.S. $= 480\cdot80 - 73\cdot41 = 407\cdot39$.

5. Summary

Following the method of estimating missing values so as to minimise the error S.S., expressions have been deduced for (i) the estimates of missing values, and (ii) the variance of the difference between any two treatments, in the case of two plots missing in any manner whatsoever in each of the designs of balanced incomplete block, Youden Square, partially balanced incomplete block and square lattice with any number of replications. All the results in the case of one missing value in these designs follow from those in the case of two missing plots and have been given in terms of the parameters of the designs. The methods of writing the equations and solving them in some particular cases for estimating the missing values when any number of plots are missing, have been indicated. The method of analysis of a partially balanced incomplete block design with two associate classes having two missing plots, has been illustrated by means of an example.

References

Cornish, E. A. .. *Ann. Eugen.*, 1940, **10**, 112–18.

——— .. *Ibid.*, 1940, **10**, 137–43.

Yates, F. .. *Emp. J. Exp. Agric.*, 1933, **1**, 129–42.

Zelen, M. .. *Biometrics*, 1954, **10**, 273–81.