MISSING PLOTS AND A RANDOMISED BLOCK DESIGN WITH BALANCED INCOMPLETENESS

By M. N. Das

Indian Council of Agricultural Research, New Delhi

1. Introduction

The problem of the analysis of data in experiments with missing plots was first dealt with by Allan and Wishart (1930) who by the method of fitting constants derived estimates of the missing values in the case of one missing plot. Subsequently Yates (1933) gave a method of estimating any number of missing observations by minimising the error sum of squares obtained by substituting unknowns for the missing observations. In the case of one missing observation he derived an expression for the standard error of the difference between two treatment means one of which has a missing replicate. For cases involving more than one missing plot, he suggested an approximation to derive the *S.E.*s for the various types of comparisons. Later Baten (1952) derived, for the case of two missing plots, formulæ for the *S.E.*s of the difference between two treatments, one or both of which have missing replicates.

An attempt has been made in the present paper to derive a general method of analysis of a randomised block design when any number of plots are missing in any manner except that there should be at least one treatment with no missing replicate. It has also been shown that if the plots are missing in a certain manner, an incomplete randomised block design with efficiency of comparison greater than that of the balanced incomplete block design can be obtained. Incomplete data where (1) the number of affected blocks and treatments are each equal to the number of missing plots, or (2) all the missing plots are in a single block, or (3) all the missing plots are the replicates of the same treatment, are some of the particular cases of the above design. As such the adjusted sum of squares due to the treatments and the *S.E.* of the difference between any two treatments, are readily obtainable in all these cases from the general results derived in this paper.

An exhaustive set of formulæ giving the *S.E.*s of the different types of contrasts has been given in the case of three and four plots missing in any manner whatsoever.

2. General Method of Analysis

In a randomised block experiment with k treatments and r replications let x plots be missing such that n_j plots are missing in the jth block and q_i replicates are missing for the ith treatment, so that $\Sigma n_j = \Sigma q_i = x$.

If t_i represents the effect of the ith treatment and b_j that of the jth block then by the method of fitting constants by least squares it follows that the best estimates of t_i's and b_j's can be obtained from the following normal equations:—

$$T_i = (r - q_i)\, t_i + \sum_j b_j \tag{1}$$

$$B_j = (k - n_j)\, b_j + \sum_i t_i \tag{2}$$

where T_i is the total of all the available replications of the ith treatmeat, B_j is the jth block total, $\sum_j$ extends over those blocks where the ith treatment is not missing and $\sum_i$ extends over those treatments which are not missing in the jth block.

On eliminating b_j's and any one of the unaffected treatments from (1) and (2) together with the additional restriction $\Sigma\, t = 0$, it has been shown by the author (Das, 1953) that the normal equations reduce to

$$rt_i + \sum_{m \neq i} t_m \left(\sum_{j\,(m)} \frac{1}{k - n_j} \right) = Q_i, \tag{3}$$

$$(r - q_m)\, t_m + \sum_{m' \neq m} t_{m'} \left(\sum_{j\,(m')}' \frac{1}{k - n_j} \right) = Q_m, \tag{4}$$

where i stands for the unaffected treatments, *i.e.*, those having no replicate missing; m, for the affected ones; $\sum_{j(m)}$ extends over those blocks where the mth treatment is missing; $\sum_{j\,(m')}'$ extends over those blocks where the m'th treatment is missing but not the mth one; Q_i and Q_m are the adjusted totals for the ith and mth treatments respectively. Actually, $Q_i = T_i - \Sigma \bar{y}_j$ where $\bar{y}_j$ is the mean of the jth block and the summation Σ extends over all the blocks; and

$$Q_m = T_m - \Sigma\, \bar{y}_j + \sum_{j\,(m)} \bar{y}_j.$$

As equations (4) contain the effects of the affected treatments only, they can be solved independently of the other treatments which can easily be obtained from (3) once the t_m's are known. The number of constants involved for solution is thus equal to that of the affected treatments.

The equations (3) and (4) can be written in a different form also. If

$$S_j = \frac{\Sigma\, t_{m(j)}}{k - n_j},$$

where $\Sigma\, t_{m(j)}$ denotes the sum of the effects of those treatments which are missing in the jth block, then (3) and (4) can be written as

$$rt_i + \Sigma\, S_j = Q_i, \tag{5}$$

$$(r - q_m)\, t_m + \sum_j{}' S_j = Q_m \tag{6}$$

where Σ denotes summation over all blocks and $\sum_j{}'$ is the sum over those blocks where the mth treatment is not missing.

Dividing (6) by $(r - q_m)$ and summing over the treatments missing in the jth block, we get

$$(k - n_j)\, S_j + \sum_{m(j)} \frac{\sum_j{}' S_j}{r - q_m} = \sum_{m(j)} \frac{\theta_m}{r - q_m}\,. \tag{7}$$

Thus summing over the treatments missing in the other blocks, there will be as many equations in S_j as the number of blocks affected. Once the S_j's are known from these equations t_i and t_m can be obtained from equations (5) and (6). So, if the number of affected blocks be less than that of the affected treatments, equations in S_j should be solved.

The equations in both (4) and (7) can be solved by iterative method. In (4), $Q_m/(r - q_m)$ may be taken as a good approximation for t_m and in (7),

$$\left\{\sum_{m(j)} \frac{Q_m}{r - q_m}\right\}\Big/ (k - n_j),$$

can be taken to be an approximate value of S_j.

After the treatment effects have been estimated the sum of squares due to the treatments adjusted for the block effects can be obtained from ΣtQ, the summation extending over all the treatments. The error sum of squares can be obtained by subtracting ΣtQ from 'within block sum of squares', ignoring the treatment classification. The degrees of freedom for the error sum of squares are $(r-1)(k-1)-x$.

The variance of the difference between any two unaffected treatments is $2\sigma^2/r$. If one or both of the treatments are affected the variance of the difference may be obtained by first substituting $+1$ and -1 for the Q's corresponding to the treatments and zero for the other Q's in the normal equations and then solving for the two treatments. If t_i and t_m are the two treatments and their solutions are t_i' and t_m', then the variance of $(t_i - t_m)$ is $\sigma^2(t_i' - t_m')$.

In the case when both the treatments are affected, the variance can be found easily if they are balanced. Two treatments may be said to be balanced when an equal number of replicates of each of them is missing, there being either no other missing plots in the block or blocks in which the treatments are missing or if there be any they should be the replicates of the same set of other treatments each of which should be missing in each of the blocks where the two treatments are missing and in none other. The variance of the difference of any two balanced treatments can be obtained by subtracting the equations in S_j corresponding to the two treatments, substituting in terms of t_m's for the S_j's remaining after the subtraction and then collecting the coefficients of the two treatments. The coefficients of both the treatments will be the same but for the sign and the other t's will get cancelled. If c be this coefficient, then the variance is $2\sigma^2/c$.

In the general case of any number of missing plots, an algebraic solution of the normal equations is not possible, but the number of unknowns can be reduced so as always to be less than those involved in the case of analysis by minimising the error sum of squares, where the number of unknowns is the same as the number of plots missing. When no two plots in the same block and no two replicates of the same treatments, are missing, the number of unknowns involved in the method described here is equal to the number of missing plots; but as will be shown in the next Section, this case falls under a special case where the algebraic solution is available.

3. A Special Case

If each of the affected blocks has a constant number of plots missing and each affected treatment a constant number of replicates missing, such that the number of times any pair of treatments is missing in the same block is also constant then the normal equations become algebraically solvable.

Thus, if

$$\left.\begin{aligned} p &= \text{the number of affected treatments such that } p < k; \\ q &= \text{the number of replicates missing per treatment;} \\ s &= \text{the number of blocks affected so that } r \geqslant s \geqslant q; \\ n &= \text{the number of plots missing per block;} \\ \lambda &= \text{the number of times that any pair of treatments is} \\ &\quad\ \text{missing in the same block, so that } 0 \leqslant \lambda \leqslant q; \end{aligned}\right\} \quad (A)$$

then

$$pq = sn; \quad \lambda\,(p-1) = q\,(n-1)$$

and the normal equations become:

$$rt_i + \frac{q\,\Sigma t_m}{k-n} = Q_i \qquad (8)$$

$$(r-q)\,t_m + \frac{(q-\lambda)}{k-n} \sum_{m' \neq m} t_{m'} = Q_m \qquad (9)$$

Solving the equations (8) and (9) and putting $v = (r-q)\,(k-n)$, we obtain the solutions:

$$\left.\begin{aligned} t_i &= \frac{Q_i}{r} - \frac{q}{r} \cdot \frac{\Sigma\, Q_m}{v + (p-1)(q-\lambda)} \\ t_m &= \frac{k-n}{v-q+\lambda} \left\{ Q_m - \frac{(q-\lambda)\,\Sigma\, Q_m}{v + (p-1)\,(q-\lambda)} \right\} \end{aligned}\right\} \qquad (10)$$

where ΣQ_m denotes the sum of the Q's corresponding to all the missing treatments.

The treatment sum of squares eliminating the block effects is thus given by

$$\frac{\Sigma Q_i^2}{r} + \frac{k-n}{v-q+\lambda}\,\Sigma Q_m^2 + \frac{(\Sigma\, Q_m)^2}{v+(p-1)\,(q-\lambda)} \times \left\{ \frac{q}{r} - \frac{(k-n)\,(q-\lambda)}{v-q+\lambda} \right\} \qquad (11)$$

The variance of the difference between two treatments, of which

$$\left.\begin{array}{l} (a) \text{ none is affected } = \dfrac{2\sigma^2}{r} \\ (b) \text{ only one is affected} \\ \quad = \sigma^2 \left[\dfrac{1}{r} + \dfrac{1}{v + (p-1)(q-\lambda)} \left\{\dfrac{q}{r} + \dfrac{k-n}{v-q+\lambda} \right.\right. \\ \qquad \left.\left. \times \overline{v + (q-\lambda)(p-2)} \right\}\right] \\ (c) \text{ both are affected } = \dfrac{2\sigma^2(k-n)}{v-q+\lambda}\,. \end{array}\right\} \quad (12)$$

4. An Incomplete Randomised Block Design with Balanced Incompleteness

If from a randomised block design, any number of plots are missing in any manner, the resulting design may be said to be the general incomplete randomised block design. The well-known incomplete block designs are actually particular cases of this design, and may be said to be perfectly incomplete in the sense that plots are missing in every block affecting every treatment, such that the remaining data are balanced or partially balanced.

In this paper, we have derived in Section 3 the method of analysis of a design which is partially incomplete such that the incompleteness, *i.e.*, the missing data are balanced so as to satisfy relations (A). The incompleteness is partial in the sense that there is at least one treatment which is not affected, *i.e.*, occurs in every block, and there may or may not be some blocks having all the treatments.

From any randomised block design with k treatments and r replications, an incomplete randomised block design with partial but balanced incompleteness can always be obtained by rejecting n plots from each of $s \leqslant r$ blocks and q replications from each of $p < k$ treatments such that the relations (A) are satisfied. When $s = r$ and $p = k$, it becomes a balanced incomplete block design.

Corresponding to any balanced incomplete block (B.I.B.) design, a series of designs with partial but balanced incompleteness can be obtained by adding $\alpha > 0$ more treatments and putting them all in each block, and $\beta \geqslant 0$ more blocks, each containing all the treatments. For if v', b', r', k' and λ' be the parameters of the B.I.B. design, it may be considered to be a $v' \times b'$ randomised block design with $v' - k'$ plots missing per block, $b' - r'$ replicates missing per treatment such that the number of times any pair of treatments is missing in the same block is also constant, being equal to $b' + \lambda' - 2r'$. This shows that the incompleteness of any B.I.B. design is balanced but not partial. If now α more treatments and β more blocks are added as stated above, the resulting design will have partial but balanced incompleteness such that

the number of blocks $(r) = b' + \beta$.

no. of treatments $(k) = v' + \alpha$

no. of blocks affected $(s) = b'$.

no. of plots missing per block $(n) = v' - k'$.

no. of treatments affected $(p) = v'$.

no. of replicates missing per treatment $(q) = b' - r'$

no. of times each pair of affected treatments is missing in the same block $(\lambda) = b' + \lambda' - 2r'$.

The efficiency of the comparison of any two of the v' affected treatments is, from 12 (*c*),

$$\frac{v'\lambda' + r'\alpha + \beta(k' + \alpha)}{r'k' + r'\alpha + \beta(k' + \alpha)}.$$

The efficiency of such comparisons cannot, of course, be compared with that of the B.I.B. design from which this design is obtained, as the number of replications in this case is $r' + \beta$ while in the case of the B.I.B. design, it is r' only. But if β is taken to be zero, the replications of the affected treatments become the same in both the designs. The efficiency of the comparison of affected treatments in this case, reduces to

$$\frac{v'\lambda' + r'\alpha}{r'k' + r'\alpha}$$

which is evidently greater than the efficiently of the original B.I.B. design.

5. Particular Cases

The different experiments where (*a*) the number of missing plots is equal each to number of affected blocks and the number of affected treatments, or (*b*) all the missing plots are in the same block or in the same treatment, or (*c*) a set of treatments is missing in each of several blocks, are some of the particular cases of the design with balanced incompleteness. In (*a*) the parameters of the design are given by

$p = s = x$, $n = q = 1$, $\lambda = 0$. The parameters for other cases can be easily obtained on consulting (A) in Section 3.

The incompleteness resulting from one or two missing plots is always balanced notwithstanding the manner in which the plots may be missing. The parameters corresponding to the different ways in which two plots can be missing, can be obtained from either $p = s = 2$ or $s = q = 2$ or $p = n = 2$. Baten (1952) has obtained the formulæ for the *S.E.*s of the different types of comparisons in all these three cases. His results agree with those obtained here.

The incompleteness due to three or more missing plots is not always balanced. In the case of three missing plots, out of six ways in which the plots can be missing with at least one treatment unaffected, the incompleteness is balanced in three cases. In the case of four missing plots, five out of sixteen cases have balanced incompleteness. To make the solution in the cases of three and four missing plots exhaustive, the formulæ giving the *S.E.*s of the different comparisons, between treatments in all cases of non-balanced incompleteness have been given below separately for each case. In cases of balanced incompleteness, the values of the parameters only have been given, so that the *S.E.*s can be obtained directly from the results presented in Section 3.

Formulæ for the variance of different treatment contrasts when three plots are missing

Case I.—Three blocks and three treatments are affected (balanced).

$$(p = s = 3, n = q = 1, \lambda = 0).$$

Case II.—Three blocks and two treatments are affected (non-balanced). (Affected treatments are denoted by t_1 and t_2. Two replications of t_1 and one of t_2 are missing.)

$$\text{Var}\,(t_1 - t_2) = \sigma^2 \frac{(k-1)\,(2v-k+4)}{v\,(v-k+1)-2} \quad \text{where } v = (r-1)\,(k-1)$$

$$\text{Var}\,(t_i - t_1) = \frac{\sigma^2}{r}\left\{2 + \frac{2vk}{v\,(v-k+1)-2}\right\} \quad \text{where } t_i \text{ stands for unaffected treatments.}$$

$$\text{Var}\,(t_i - t_2) = \frac{\sigma^2}{r}\left\{2 + \frac{k\,(v-k+1)}{v\,(v-k+1)-2}\right\}.$$

Case III.—Three blocks and one treatment are affected (balanced).

$$(p = n = 1, s = q = 3, \lambda = 0).$$

Case IV.—Two blocks and three treatments are affected (non-balanced). (t_1 and t_2 are missing in the same block and t_3 in another.)

$$\text{Var}(t_1 - t_2) = \frac{2\sigma^2}{r-1}.$$

$$\text{Var}(t_1 - t_3) = \text{var}(t_2 - t_3) = \sigma^2 \left[\frac{2v(v-r+2) - r}{(r-1)\{v(v-r+1) - 2\}}\right]$$

$$\text{Var}(t_i - t_1) = \text{var}(t_i - t_2) = \sigma^2$$

$$\times \left\{\frac{1}{r} + \frac{(vr-1)(v-r+2) - 3(r-1)}{r(r-1)\{v(v-r+1) - 2\}}\right\}$$

$$\text{Var}(t_i - t_3) = \frac{\sigma^2}{r}\left\{1 + \frac{(v+1)(v-r-1) + vk}{v(v-r+1) - 2}\right\}.$$

Case V.—Two blocks and two treatments are affected (non-balanced) (Two replications of t_1 and one of t_2 are missing.)

$$\text{Var}(t_1 - t_2) = \sigma^2 \left\{\frac{1}{r-1} + \frac{1}{r-2} + \frac{1}{v(r-2)}\right\}$$

$$\text{Var}(t_i - t_1) = \sigma^2 \left\{\frac{1}{r} + \frac{1}{r-2} + \frac{2v - r}{rv(r-2)(k-2)}\right\}$$

$$\text{Var}(t_i - t_2) = \sigma^2 \left\{\frac{1}{r} + \frac{1}{r-1} + \frac{1}{r(v-r+1)}\right\}.$$

Case VI.—One block and three treatments are affected (balanced).

$$(p = n = 3,\ s = q = 1,\ \lambda = 1).$$

Formulæ for the variance of different treatment contrasts when four plots are missing

Case I.—Four blocks and four treatments are affected (balanced).

$$(p = s = 4,\ q = n = 1,\ \lambda = 0).$$

Case II.—Four blocks and three treatments are affected (non-balanced).

(Two replications of t_1 and one each of t_2 and t_3 are missing.)

$$\text{Var}(t_1 - t_2) = \text{var}(t_1 - t_3)$$

$$= \sigma^2 \left[\frac{(k-1)\{2(v-1)(v+3) - vk\}}{(v-1)\{(v+1)(v-k+1) - 4\}}\right]$$

$$\text{Var}\,(t_2 - t_3) = \frac{2\,(k-1)}{v-1}\,\sigma^2.$$

$$\text{Var}\,(t_i - t_1) = \frac{\sigma^2}{r}\left\{1 + \frac{(v-1)\,(v+3) + k\,(v+1)}{(v+1)\,(v-k+1) - 4}\right\},$$

$$\text{Var}\,(t_i - t_2) = \text{var}\,(t_i - t_3)$$

$$= \frac{\sigma^2}{r}\left[1 + \frac{(v-1)^2\,(v+3) - k\,(k-1)\,(v-1) - k^2}{(v-1)\,\{(v+1)\,(v-k+1) - 4\}}\right]$$

Case III.—Four blocks and two treatments are affected.

(i) Three replications of t_1 and one of t_2 are missing (non-balanced).

$$\text{Var}\,(t_1 - t_2) = \frac{2\sigma^2\,(k-1)\,(v-k+3)}{v\,(v-2k+2) - 3}$$

$$\text{Var}\,(t_i - t_1) = \frac{\sigma^2}{r}\left\{1 + \frac{(v-1)\,(v+3) + vk}{v\,(v-2k+2) - 3}\right\},$$

$$\text{Var}\,(t_i - t_2) = \frac{\sigma^2}{r}\left\{2 + \frac{k\,(v-2k+2)}{v\,(v-2k+2) - 3}\right\}.$$

(ii) Two replications of each of t_1 and t_2 are missing (balanced).

$$(p = q = 2,\ s = 4,\ n = 1,\ \lambda = 0).$$

Case IV.—Four blocks and one treatment are affected (balanced).

$$(p = n = 1,\ s = q = 4,\ \lambda = 0).$$

Case V.—Three blocks and four treatments are affected (non-balanced). (t_1 and t_2 are missing in the same block and t_3 and t_4 in some other blocks.)

$$\text{Var}\,(t_1 - t_2) = \frac{2\sigma^2}{r-1}.$$

$$\text{Var}\,(t_3 - t_4) = \frac{2\,(k-1)}{v-1}\,\sigma^2.$$

$$\left.\begin{aligned}&\text{Var}\,(t_1 - t_3) = \text{var}\,(t_1 - t_4)\\ &= \text{Var}\,(t_2 - t_3) = \text{var}\,(t_2 - t_4)\end{aligned}\right\}$$

$$= \sigma^2\left[\frac{2v\,(v-1)\,(v+3) - r\,(2v^2 + v - 2)}{(r-1)\,(v-1)\,\{(v+1)\,(v-r+1) - 4\}}\right].$$

$$\text{Var}(t_i - t_1) = \text{var}(t_i - t_2)$$

$$= \sigma^2 \left[\frac{1}{r} + \frac{1}{r-1} + \frac{vr - v + r + 1}{r(r-1)\{(v+1)(v-r+1) - 4\}}\right]$$

$$\text{Var}(t_i - t_3) = \text{var}(t_i - t_4)$$

$$= \frac{\sigma^2}{r}\left[1 + \frac{(v-r+1)(v^2+vk-1) - 2(2v+k-2)}{(v-1)\{(v+1)(v-r+1) - 4\}}\right].$$

Case VI.—Three blocks and three treatments are affected (non-balanced).

(i) Two replications of t_1 and one each of t_2 and t_3 are missing. t_1 and t_2 are missing in the same block.

$$\text{Var}(t_1 - t_2) = \frac{\sigma^2(k-1)(2v^2 + 7v - 2vr - vk - 2r)}{(v-1)\{(v+1)(v-r-k+3) - 2\}}$$

$$\text{Var}(t_1 - t_3) = \sigma^2\left[\frac{(k-2)(2v^2+2v-vk) - 2(3k-4)}{(v-1)\{(v+1)(v-r-k+3) - 2\}}\right]$$

$$\text{Var}(t_2 - t_3) = \sigma^2\left[\frac{(v-k+1)\{2(k-2)(v+1) + 1\} - 3k + 5}{(v-1)\{(v+1)(v-r-k+3) - 2\}}\right]$$

$$\text{Var}(t_i - t_1) = \frac{\sigma^2}{r}$$

$$\times\left[1 + \frac{(v+k-1)(v^2 - vr + v - 1) + v(2v - r - 2) + r}{(v-1)\{(v+1)(v-r-k+3) - 2\}}\right]$$

$$\text{Var}(t_i - t_2) = \frac{\sigma^2}{r}$$

$$\times\left[1 + \frac{v(v+k-1)(v-r-k+3) + (v-1)(v-r-k+1)}{(v-1)\{(v+1)(v-r-k+3) - 2\}}\right]$$

$$\text{Var}(t_i - t_3) = \frac{\sigma^2}{r}$$

$$\times\left[1 + \frac{(v-k+1)(v^2 - vr + vk - 1) - r(2k-3)}{(v-1)\{(v+1)(v-r-k+3) - 2\}}\right].$$

(ii) Two replications of t_1 and one each of t_2 and t_3 are missing. t_2 and t_3 are missing in the same block.

$$\text{Var}(t_1 - t_2) = \text{var}(t_1 - t_3)$$

$$= \frac{\sigma^2}{r-1}\left[1 + \frac{(v+2)(v-r+2)}{v(r-2)(k-2)-4}\right]$$

$$\text{Var}(t_2 - t_3) = \frac{2\sigma^2}{r-1}.$$

$$\text{Var}(t_i - t_1) = \frac{\sigma^2}{r}\left\{1 + \frac{v(v-r+k+1) - 2(r+1)}{v(r-2)(k-2)-4}\right\}.$$

$$\text{Var}(t_i - t_2) = \text{var}(t_i - t_3)$$

$$= \sigma^2\left[\frac{1}{r} + \frac{1}{r-1} + \frac{v(r-2)+2}{r(r-1)\{v(r-2)(k-2)-4\}}\right].$$

Case VII.—Three blocks and two treatments are affected (non-balanced).

(i) Two replications of each of t_1 and t_2 are missing.

$$\text{Var}(t_1 - t_2) = \frac{2(k-1)}{v-1}\sigma^2.$$

$$\text{Var}(t_i - t_1) = \text{V}(t_i - t_2) = \sigma^2\left\{\frac{1}{r} + \frac{1}{r(v_1+1)}\right.$$

$$\left. + \frac{k-1}{r(k-2)(v_1+1)} + \frac{v_1(k-1)}{v_1^2 - 1}\right\}$$

where $v_1 = (r-2)(k-1)$.

(ii) Three replications of t_1 and one of t_2 are missing.

$$\text{Var}(t_1 - t_2) = \sigma^2\left\{\frac{1}{r-1} + \frac{1}{r-3} + \frac{2}{v(r-3)}\right\}.$$

$$\text{Var}(t_i - t_1) = \sigma^2\left\{\frac{1}{r} + \frac{1}{r-3} + \frac{(r-1)(3k-5)-2}{rv(r-3)(k-2)}\right\}$$

$$\text{Var}(t_i - t_2) = \sigma^2\left\{\frac{1}{r} + \frac{1}{r-1} + \frac{1}{v-r+1}\right\}.$$

Case VIII.—Two blocks and four treatments are affected (non-balanced).

(i) t_1 and t_2 are missing in the same block and t_3 and t_4 in another.

$$\text{Var}(t_1 - t_2) = \text{var}(t_3 - t_4) = \frac{2\sigma^2}{r-1}.$$

$$\left.\begin{array}{l}\text{Var}(t_1 - t_3) = \text{var}(t_1 - t_4) \\ = \text{Var}(t_2 - t_3) = \text{var}(t_2 - t_4)\end{array}\right\} = \frac{2\sigma^2}{r-1}\left(1 + \frac{1}{\nu_1 - 2}\right)$$

where $\nu_1 = (r-1)(k-2)$.

$$\text{Var}(t_i - t_m) = \sigma^2\left\{\frac{1}{r} + \frac{1}{r(\nu_1 + 2)} + \frac{\nu_1^2 - 2}{(r-1)(\nu_1^2 - 4)}\right\}$$

where t_m stands for affected treatments.

(ii) t_1 is missing in one block and the rest in another.

$$\text{Var}(t_1 - t_2) = \text{var}(t_1 - t_3) = \text{var}(t_1 - t_4)$$

$$= \frac{\sigma^2}{r-1}\left\{1 + \frac{(\nu+1)(\nu - 2r + 3)}{\nu(\nu - 2r + 2) - 3}\right\}.$$

$$\text{Var}(t_i - t_1)$$

$$= \frac{\sigma^2}{r}\left\{2 + \frac{\nu - 2r + 2 + \nu(k-3)}{\nu(\nu - 2r + 2) - 3}\right\}.$$

$$\text{Var}(t_i - t_2) = \text{var}(t_i - t_3) = \text{var}(t_i - t_4)$$

$$= \sigma^2\left\{\frac{1}{r} + \frac{1}{r-1} + \frac{\nu r - \nu + 1}{\nu(\nu - 2r + 2) - 3}\right\}.$$

Case IX.—Two blocks and three treatments are affected (non-balanced). Two replications of t_1 and one each of t_2 and t_3 are missing.

(i) Two plots are missing in each block.

$$\text{Var}(t_1 - t_2) = \text{var}(t_1 - t_3)$$

$$= \frac{\sigma^2}{r-2}\left\{1 + \frac{\nu_1(\nu_2 + 1) - 1}{\nu_1^2 - 1}\right\}$$

where $\nu_1 = (r-1)(k-2)$ and $\nu_2 = (r-2)(k-2)$

$$\text{Var}(t_2 - t_3)$$

$$= \frac{2(k-2)}{\nu_1 - 1}\sigma^2.$$

$\text{Var}\,(t_i - t_1)$

$$= \sigma^2 \left\{\frac{1}{r} + \frac{1}{r-2} + \frac{2\,(r-1)}{r\,(r-2)\,(v_1+1)}\right\}.$$

$\text{Var}\,(t_i - t_2) = \text{var}\,(t_i - t_3)$

$$= \frac{\sigma^2}{r}\left\{2 + \frac{v_1\,(k-1)}{v_1^2 - 1}\right\}.$$

(ii) Three plots are missing in one block and one in another.

$\text{Var}\,(t_1 - t_2) = \text{V}\,(t_1 - t_3)$

$$= \sigma^2 \left\{\frac{1}{r-1} + \frac{1}{r-2} + \frac{1}{v\,(r-2)}\right\}.$$

$$\text{Var}\,(t_2 - t_3) = \frac{2\sigma^2}{r-1}.$$

$\text{Var}\,(t_i - t_1)$

$$= \sigma^2 \left\{\frac{1}{r} + \frac{1}{r-2} + \frac{2\,(v-r)}{vr\,(r-2)\,(k-3)}\right\}$$

$\text{Var}\,(t_i - t_2) = \text{var}\,(t_i - t_3)$

$$= \sigma^2 \left\{\frac{1}{r} + \frac{1}{r-1} + \frac{1}{r\,(r-1)\,(k-3)}\right\}.$$

Case X.—Two blocks and two treatments are affected (balanced).

$$(p = q = s = n = \lambda = 2).$$

Case XI.—One block and four treatments are affected (balanced).

$$(p = n = 4,\ q = s = 1,\ \lambda = 1).$$

6. An Example to Illustrate the General Method

The general method of analysis when any number of plots are missing in a randomised block layout in any manner, except that there is at least one treatment unaffected, has been illustrated in the following example. The data used are the same as those taken by Yates (1933). The table given below shows the cell-frequencies of the available data together with the block averages and treatment totals both adjusted and unadjusted.

TABLE I

Cell-frequencies, block totals and treatment totals, both adjusted and unadjusted

Blocks	Treatments								Totals B_j	Averages $\bar{y}_j$
	O (1)	N (2)	K (3)	P (4)	NK (5)	NP (6)	PK (7)	NPK (8)		
1	1	1	1	1	0	1	1	1	20·48 (7)	2·926
2	1	1	1	1	1	1	1	1	25·33 (8)	3·167
3	0	1	1	1	1	1	1	1	19·38 (7)	2·769
4	1	1	1	1	1	1	1	1	21·81 (8)	2·726
5	1	1	1	1	1	1	1	0	25·08 (7)	3·583
6	1	1	1	1	1	1	0	0	21·92 (6)	3·653
7	1	0	1	1	1	0	1	1	22·39 (6)	3·732
8	1	1	1	0	1	0	1	1	19·10 (6)	3·183
9	1	1	1	1	1	1	1	1	22·59 (8)	2·824
10	1	1	1	1	1	1	1	1	25·77 (8)	3·221
Totals T_i	27·91 (9)	24·96 (9)	33·41 (10)	33·99 (9)	28·52 (9)	24·37 (8)	25·50 (9)	25·59 (8)		
Adjusted totals (Q_i)	−1·505	−3·092	1·626	5·389	−·329	−·499	−2·631	1·042		

The condition of applicability of the method is satisfied as treatment No. 3 is present in every block. As the number of blocks affected is less than that of the affected treatments it will be advantageous to solve the equations in s_j, but for the sake of illustration the equations have also been written in terms of t_i's, *viz.*,

$$10t_3 + \frac{t_1}{7} + \frac{t_2}{6} + \frac{t_4}{6} + \frac{t_5}{7} + \frac{2t_6}{6} + \frac{t_7}{6} + t_8\left(\frac{1}{7} + \frac{1}{6}\right) = 1{\cdot}626$$

$$9t_1 + \frac{t_2}{6} + \frac{t_4}{6} + \frac{t_5}{7} + \frac{2t_6}{6} + \frac{t_7}{6} + t_8\left(\frac{1}{7} + \frac{1}{6}\right) = -1{\cdot}505$$

$$9t_2 + \frac{t_1}{7} + \frac{t_4}{6} + \frac{t_5}{7} + \frac{t_6}{6} + \frac{t_7}{6} + t_8\left(\frac{1}{7} + \frac{1}{6}\right) = -3{\cdot}092$$

$$9t_4 + \frac{t_1}{7} + \frac{t_2}{6} + \frac{t_5}{7} + \frac{t_6}{6} + \frac{t_7}{6} + t_8\left(\frac{1}{7} + \frac{1}{6}\right) = 5{\cdot}389$$

$$9t_5 + \frac{t_1}{7} + \frac{t_2}{6} + \frac{t_4}{6} + \frac{2t_6}{6} + \frac{t_7}{6} + t_8\left(\frac{1}{7} + \frac{1}{6}\right) = -0\cdot329$$

$$8t_6 + \frac{t_1}{7} + \frac{t_5}{7} + \frac{t_7}{6} + t_8\left(\frac{1}{7} + \frac{1}{6}\right) = -0\cdot499$$

$$9t_7 + \frac{t_1}{7} + \frac{t_2}{6} + \frac{t_4}{6} + \frac{t_5}{7} + \frac{2t_6}{6} + \frac{t_7}{6} + \frac{t_8}{7} = -2\cdot631$$

$$8t_8 + \frac{t_1}{7} + \frac{t_2}{6} + \frac{t_4}{6} + \frac{t_5}{7} + \frac{2t_6}{6} = 1\cdot042$$

Plots are missing in block numbers 1, 3, 5, 6, 7 and 8 and the corresponding s_i's are

$$s_1 = \frac{t_5}{7},\ s_3 = \frac{t_1}{7},\ s_5 = \frac{t_8}{7},\ s_6 = \frac{t_7 + t_8}{6},\ s_7 = \frac{t_2 + t_6}{6},\ s_8 = \frac{t_4 + t_6}{6}$$

Now the normal equations can be written as

$$10t_3 + s_1 + s_3 + s_5 + s_6 + s_7 + s_8 = 1\cdot626 \quad (1)$$

$$9t_1 + s_1 + s_5 + s_6 + s_7 + s_8 = -1\cdot505 \quad (2)$$

$$9t_2 + s_1 + s_3 + s_5 + s_6 + s_8 = -3\cdot092 \quad (3)$$

$$9t_4 + s_1 + s_3 + s_5 + s_6 + s_7 = 5\cdot389 \quad (4)$$

$$9t_5 + s_3 + s_5 + s_6 + s_7 + s_8 = -0\cdot329 \quad (5)$$

$$8t_6 + s_1 + s_3 + s_5 + s_6 = -0\cdot499 \quad (6)$$

$$9t_7 + s_1 + s_3 + s_5 + s_7 + s_8 = -2\cdot631 \quad (7)$$

$$8t_8 + s_1 + s_3 + s_7 + s_8 = 1\cdot042 \quad (8)$$

All these equations excepting the first which contains the treatment which has no missing replicate can be expressed in terms of s_i's. Thus from the equations (2), (5) and (8) above,

$$63\,s_3 + s_1 + s_5 + s_6 + s_7 + s_8 = -1\cdot505 \quad (9)$$

$$63\,s_1 + s_3 + s_5 + s_6 + s_7 + s_8 = -0\cdot329 \quad (10)$$

$$56s_5 + s_1 + s_3 + s_7 + s_8 = 1\cdot042 \quad (11)$$

Dividing equations (7) and (8) by 9 and 8 respectively and then adding them the equations corresponding to s_6 will be obtained. Two more equations corresponding to s_7 and s_8 can be obtained similarly from equations (3) & (6) and (4) & (6) respectively. Getting rid of the denominators by suitable multiplication these become

$$432\, s_6 + 17\,(s_1 + s_3 + s_7 + s_8) + 8 s_5 = -\,11{\cdot}670 \qquad (12)$$

$$432\, s_7 + 17\,(s_1 + s_3 + s_5 + s_6) + 8\, s_8 = -\,29{\cdot}227 \qquad (13)$$

$$432\, s_8 + 17\,(s_1 + s_3 + s_5 + s_6) + 8\, s_7 = 38{\cdot}621 \qquad (14)$$

The equations (9) to (14) have been solved by iterative method. The first approximation obtained by dividing $-\,1{\cdot}505$ by 63 and equating the quotient to s_3 and so on, are found to give solution correct upto the third decimal place. The solution correct upto the fourth place are as below:—

s_1	s_3	s_5	s_6	s_7	s_8
$-{\cdot}0050$	$-{\cdot}0241$	$+{\cdot}0166$	$-{\cdot}0271$	$-{\cdot}0678$	$+{\cdot}0922$

Substituting these values in equations (1) to (8) the treatment effects are as follows:—

t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
$-{\cdot}1682$	$-{\cdot}3494$	${\cdot}1641$	${\cdot}6107$	$-{\cdot}0354$	$-{\cdot}0574$	$-{\cdot}2937$	${\cdot}1308$

The *s.s.* due to the treatments obtained from ΣtQ is $5{\cdot}8407$.

The *s.s.* within blocks $= 32{\cdot}1012 - 8{\cdot}5690 = 23{\cdot}5322$.

Hence the Error *s.s.*

$$= 23{\cdot}5322 - 5{\cdot}8407$$

$$= 17{\cdot}6915 \text{ for 54 degrees of freedom.}$$

This agrees with the sums of squares obtained by Yates, *viz.*, $5{\cdot}8420$ and $17{\cdot}6902$ for treatments and error respectively.

There are two pairs of balanced treatments, *viz.*, t_1 and t_5; and t_2 and t_4.

By subtracting the equations corresponding to t_2 and t_4 and substituting in terms of t's for the remaining s_j's, we obtain:

$$\frac{53}{6}(t_2 - t_4) = Q_2 - Q_4.$$

Hence

$$V(t_2 - t_4) = \frac{12}{53}\sigma^2.$$

Similarly,

$$V(t_1 - t_5) = \frac{7}{31}\sigma^2.$$

7. Another example to illustrate the method of Analysis of a design with balanced incompleteness.

The data analysed are taken from an experiment in randomised block design with 7 replicates, conducted for selecting the best varieties from among 8 varieties of Mung. Actually the data were complete, but for the sake of illustrations 12 plot yields have been deliberately omitted, 2 from each of 6 blocks, so that four treatments were affected and the incompleteness was balanced. As in the previous example, the table below shows the cell-frequencies of the remaining data together with the block averages and treatment totals, both adjusted and unadjusted.

TABLE II

Cell-frequencies, block totals and treatment totals, both adjusted and unadjusted

Treatments / Blocks	Number of the different selections of Mung								Totals	Averages
	1	2	3	4	5	6	7	8		
1	0	0	1	1	1	1	1	1	113·5 (6)	18·92
2	0	1	0	1	1	1	1	1	108·0 (6)	18·00
3	0	1	1	0	1	1	1	1	80·5 (6)	13·42
4	1	0	0	1	1	1	1	1	132·5 (6)	22·08
5	1	0	1	0	1	1	1	1	134·0 (6)	22·33
6	1	1	0	0	1	1	1	1	117·5 (6)	19·58
7	1	1	1	1	1	1	1	1	138·0 (8)	17·25
Totals ..	94·5 (4)	72·5 (4)	79·5 (4)	62·0 (4)	139·5 (7)	166·0 (7)	109·5 (7)	100·5 (7)	824·0 (44)	
Adjusted totals (Q_i)	13·25	4·25	7·58	−14·25	7·92	34·42	−22·08	−31·08		

The incompleteness is balanced with parameters $p = 4$, $q = 3$, $s = 6$, $n = 2$, $\lambda = 1$.

Hence the treatment effects can be obtained from the equations:—

$$t_m = \frac{k-n}{v-q+\lambda} \left\{ Q_m - \frac{(q-\lambda)\,\Sigma\, Q_m}{v+(p-1)(q-\lambda)} \right\}$$

$$= \frac{3}{11} \cdot Q_m - 0\cdot 1970 \; (m = 1, 2, 3, 4)$$

and

$$t_i = \frac{Q_i}{r} - \frac{q}{r} \cdot \frac{\Sigma\, Q_m}{v+(p-1)(q-\lambda)} = \frac{Q_i}{7} - 0\cdot 1548$$

$$(i = 5, 6, 7, 8)$$

The solutions for t's are

t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
3·4167	0·9621	1·8712	−4·0834	0·9762	4·7619	−3·3096	−4·5953

Thus

Adjusted *s.s.* due to treatments ΣtQ = 509·28

Within Blocks *s.s.* = 1067·17

Error *s.s.* = 1067·17 − 509·28 = 557·89

for 30 degrees of freedom.

$$\text{Var}\,(t_m - t_m') = \frac{6}{11}\,\sigma^2$$

and

$$\text{Var}\,(t_i - t_m) = \frac{317}{770}\,\sigma^2.$$

8. Summary

A method of analysis of a randomised block design with any number of plots missing in any manner, has been described. An incomplete randomised block design with partial but balanced incompleteness and with efficiency greater than that of a balanced incomplete block design, has been obtained. Expressions for variances of the difference between treatments have been deduced in the case of three and four plots missing in any manner whatsoever. The methods have been illustrated by means of two examples.

References

Allan, F. E. and Wishart, J. *J. Agric. Sci.*, 1930, **20**, 399–406.

Baten, W. D. .. *Biometrics*, 1952, **8**, 42–50.

Das, M. N. .. *J. Ind. Sec. Agric. Stat.*, 1953, **5**, 161–178.

Yates, F. .. *Emp. J. Exp. Agric.*, 1933, **1**, 129–142.