

LOSS OF INFORMATION IN INCOMPLETE BLOCK DESIGNS

M. V. R. PRASADA RAO

Sri Venkateswara College, Delhi University, Delhi

(Received : April, 1987)

SUMMARY

A generalised model for analysis of variance is introduced. Methods for estimation of parameters in this model for non-orthogonal data are provided. The results are applied for finding the total loss of information of balanced and efficiency balanced designs without any restriction on n_{ij} 's, the elements of the incidence matrix of a design.

Keywords: Non-orthogonal data; Eigen value; Contrast matrix, balanced and efficiency balanced design; Total loss of information.

Introduction

The model used for comparing treatment effects using the technique of analysis of variance in two-way classified data is:

$$Y_{ijk} = \mu + t_i + b_j + e_{ijk}$$
$$(i = 1, 2, \dots, v; j = 1, 2, \dots, r$$
$$k = 1, 2, \dots, n_{ij})$$

where Y_{ijk} is the k th observation belonging to i th class of a treatment A , having v levels and j th class of another factor, B , with r levels, $j = t_i, b_j$ etc., are the parameters representing the effects of different levels of the two factors; n_{ij} is the number of observations in the cell defined by the i th class of A and j th class of B and need not be zero or unity only.

There are many situations where the above model is not suitable. For example, in factorial experiments a different model is used. Again, in

bio-assays the interest is not to compare the dose effects directly but to estimate certain contrasts of the dose effects and as such a reparametrised model is more suitable. Accordingly, we propose the following model which can be applied in general in all situations :

$$Y_{ijk} = \mu + \sum_{m=1}^v c_{im} a_m + b_j + e_{ijk}$$

$$(i = 1, 2, \dots, v; j = 1, 2, \dots, r) \quad (1.1)$$

where c_{im} 's are known coefficients of the unknown parameters a_m which have to be estimated. The model involves the matrix of the known constants c_{ij} 's namely,

$$C = [c_{ij}]_{v \times v} \quad (1.2)$$

$$(i = 1, 2, \dots, v; j = 1, 2, \dots, v)$$

C is subject to the restriction that it is non-singular. Actually, the parameters $a_1, a_2 \dots a_v$ can be considered to be certain functions of level effects of factors. When the data are non-orthogonal (in the usual sense, the problem now is how to get the estimates of the parameters and conduct the appropriate analysis of variance.

1.1 Estimation and Analysis

Using matrix notation the normal equations for estimation of the parameters in the above model through least squares are as below :

$$C'T = C'ND + C'NCA + C'P'b \quad (1.3)$$

$$B = KE + Kb + PCA \quad (1.4)$$

where $C = [c_{ij}]_{v \times v}$; $T' = [T_1, T_2, \dots, T_v]_{1 \times v}$

$$P = \begin{bmatrix} n_{11} & n_{21} & \dots & n_{v1} \\ n_{12} & n_{22} & \dots & n_{v2} \\ \vdots & & & \\ n_{1r} & n_{2r} & \dots & n_{vr} \end{bmatrix}_{r \times v}$$

$$\begin{aligned} A' &= [a_1, a_2, \dots, a_v]_{1 \times v} \\ D' &= [m, m, \dots, m]_{1 \times v} \end{aligned}$$

$$B' = [B_1, B_2, \dots, B_r]_{1 \times r}; b' = [b_1, b_2, \dots, b_r]_{1 \times r}$$

$$E' = [m, m, \dots, m]_{1 \times r}$$

$$N = \begin{bmatrix} n_{1.} & 0 & 0 & \dots & 0 \\ 0 & n_{2.} & 0 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & \dots & n_{v.} \end{bmatrix}_{v \times v}; K = \begin{bmatrix} n_{.1} & 0 & \dots & 0 \\ 0 & n_{.2} & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & n_{.r} \end{bmatrix}_{r \times v}$$

and $n_{i.} = \sum_{j=1}^r n_{ij}$; $n_{.j} = \sum_{i=1}^v n_{ij}$, T_i is the total of the observations

belonging to the i th level of A ; B_j is the total of the observations belonging to the j th level of B ; a_i 's are the parameters to be estimated. Eliminating b from (1.3) using (1.4) we get

$$C' [T - P^1 K^{-1} B] = C' [N - P' K^{-1} P] C A \quad (1.5)$$

That is,

$$R = C' [N - P' (K^{-1} P) C A] \quad (1.6)$$

where R is the $v \times 1$ vector, $C' [T - P' K^{-1} B]$.

This is the general form of the reduced normal equations for estimating the parameters a_i 's.

Further investigation in regard to solution of these equations and analysis of variance depends on the specific nature of the data and the problem. Actually, the nature of the problem provides the C matrix. For example, if the design be factorial, C -matrix should be taken as an orthogonal matrix where rows give the co-efficients of different main-effects and interaction contrasts. In the present investigation we have used the model for obtaining the total loss of information in balanced and efficiency balanced designs where n_{ij} can be anything.

2. Investigation of Loss of Information in Incomplete Block Designs

Kshirsagar [4] and Tyagi [6] investigated the problem of finding the total loss of information in designs following Yates conjecture that the total loss of information in a confounded design is one less than the number of blocks per replication. They, however, did not consider the cases where n_{ij} , the number of observations or i th treatment in j th block is greater than unity. The present model can be very conveniently used to investigate this problem particularly when the numbers of replications are unequal and n_{ij} can be greater than or equal to unity.

2.1 Total Loss of Information in Balanced Designs

Let $P_{v \times v}$ denote the incidence matrix of a variance balanced incomplete

block design. Let r_1, r_2, \dots, r_v be the numbers of replications of the treatments and let K_1, K_2, \dots, K_b be the block sizes and n_{ij} denote the number of times the i th treatment occurs in j th block. Further we define,

$A = C^{-1}t$ where t' is the row vector (t_1, t_2, \dots, t_v) formed of the effects of treatment and A is the same $v \times 1$ vector as defined earlier.

The reduced normal equations for such designs are

$$C' [N - P' K^{-1} P] C A = R \quad (2.1)$$

If C can be an orthogonal matrix, the above equations reduce to

$$gA = R \quad (2.2)$$

where g is a scalar independent of the matrix C ; $(N - P' K^{-1} P)$ is the symmetrical variance-covariance matrix of Q_i 's where Q_i 's are adjusted treatment totals and C is any orthogonal matrix. It is well known that if the product $C' (N - P' K^{-1} P) C$ becomes a diagonal matrix, then the diagonal elements are the eigen values of $(N - P' K^{-1} P)$. In the present case each of the diagonal elements is g . Hence g is the eigen value of multiplicity $(v - 1)$. It can also be shown that the eigen value is equal to $\text{Var} (Q_i) - \text{Cov.} (Q_i, Q_j)$ for all balanced designs.

From this consideration it follows in general that if the solutions of the reduced normal equations

$$[C' (N - P' K^{-1} P) C] A = R$$

be of the form $A = G^{-1}R$, where G is a diagonal matrix, then the elements of G are the different eigen values of the variance-covariance matrix of Q_i 's. In this situation the design can be said to be balanced in the sense that the different parametric contrasts of treatment effects are estimable mutually independently though perhaps with unequal variances when the diagonal elements of G are unequal.

These results are illustrated by taking a variance balanced design given by Das and Ghosh [1]. The following is the incidence matrix of variance balanced design obtainable from the method of Das and Ghosh [1] presented in Section 5.1.

	Treatments							
I	2	3	4	5	6	7	8	
0	1	1	0	1	0	0	0	
0	0	1	1	0	1	0	0	
0	0	0	1	1	0	1	0	
0	0	0	0	1	1	0	1	
0	1	0	0	0	1	1	0	
0	0	1	0	0	0	1	1	
0	1	0	1	0	0	0	1	
5	1	1	1	1	1	1	1	

The parameters of the design are

$$V = 8, b = 8, r_1 = 5, r_2 = 4, K_1 = 3, K_2 = 12.$$

We take C as the following orthogonal matrix :

$$\frac{1}{2\sqrt{2}} \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{vmatrix}$$

Using this matrix we find that the reduced normal equations are

$$\frac{40}{12} A = R.$$

Thus $40/12$ is the eigen value with multiplicity 7.

$$A = \frac{12}{40} R$$

Variance of the estimate of any individual a_i is the coefficient of the corresponding R_i in the solution of a_i , that is

$$\text{Var}(a_i) = \frac{12}{40} \sigma^2$$

In general, variance of the estimate of any a_i is $\text{Var}(a_i) = \sigma^2/g$.

Moreover, the covariance of the estimates of any two a_i 's is zero. This ensures that the information recovered on different a_i 's is additive.

For investigation of efficiency of such a design let us compare it against a randomised block design with replications equal to the average of the replications in the above variance balanced designs. Let \bar{r} denote the average replication. Then information recovered for estimating any contrast relative to this randomised block design is

$$\frac{\sigma^2}{\bar{r}} / \frac{\sigma^2}{g} = \frac{g}{\bar{r}}$$

Loss of information is $1 - \frac{g}{\bar{r}}$ (2.3)

Hence, total loss of information is

$$(v - 1) \left(1 - \frac{g}{r} \right) \quad (2.4)$$

as there are $(v - 1)$ mutually orthogonal contrasts which are estimable mutually independently.

In the example considered above

$$g = \frac{40}{12}, v = 8 \quad \text{and} \quad \bar{r} = \frac{33}{8}$$

Hence total loss of information is

$$7 \left(1 - \frac{40}{12} \times \frac{8}{33} \right) = \frac{133}{99}$$

According to Tyagi (1969) total loss of information when the replications are unequal in a balanced design is $(b/\bar{r}) - 1$. That is, for the design

$$\text{under illustration it is } \frac{\frac{8}{33}}{8} - 1 = \frac{64}{33} - 1 = \frac{93}{99}$$

Tyagi has considered designs when a treatment can occur in a block at most once. But here we have taken the general case without this assumption. In this particular example one $n_j > 1$ ($n_{18} = 5$). Hence, the two results need not be the same.

The total loss of information presented in (2.4) holds for all balanced designs whether with or without equal numbers of replications for any type of cell frequencies. If the replications be equal and $n_j \leq 1$ then in the variance balanced designs which are the usual BIB designs, the eigen value g of the variance-covariance matrix of the adjusted totals is rE where E is the efficiency factor of the design, that is,

$$E = \frac{v\lambda}{rk}$$

$$\begin{aligned} \text{Hence, total loss of information} &= (v - 1) \left(1 - \frac{rE}{r} \right) \\ &= (v - 1) (1 - E) \\ &= \frac{b}{r} - 1 \end{aligned}$$

Thus the result obtained here agrees with the finding of Kshirsagar [4] for binary designs with equal number of replications.

3. Loss of Information in Efficiency Balanced Designs

Puri and Nigam [5], Dey and Singh [3], Das and Ghosh [1] contributed several series of efficiency balanced designs. Each of these series of designs is in one form or other a type of reinforced design (Das [2]) with usually one extra treatment and several extra blocks. It appears so far that no efficiency balanced design is available which is not of the above form. Accordingly, we shall restrict ourselves to the analysis of the above type of designs only.

Let t_0 denote the extra treatment and t_1, t_2, \dots, t_v denote the V other treatments in an efficiency balanced design. For analysing designs through the general model we shall get the set of reduced normal equations as in (2.1). For efficiency balanced designs of the above type we shall choose 'C'-matrix as shown below :

$$C = \begin{vmatrix} \frac{1}{\sqrt{v+1}} & \frac{1}{\sqrt{v+1}} & \cdots & \frac{1}{\sqrt{v+1}} \\ \frac{v}{\sqrt{v(v+1)}} & \frac{-1}{\sqrt{v(v+1)}} & \cdots & \frac{-1}{\sqrt{v(v+1)}} \\ \vdots & & & \\ 0_{v-1 \times 1} & \cdots & \cdots & M_{v-1 \times 1} \end{vmatrix}_{(v+1) \times (v+1)} \quad (3.1)$$

where $0_{1 \times v-1}$ is a row vector with each of the elements 0 and $M_{(v-1) \times (v)}$ is the last $(v-1)$ rows of a $v \times v$ orthogonal matrix in which the first row elements are all equal.

Using this type of orthogonal matrix for C' the reduced normal equations as indicated in (1.6) for efficiency balanced designs come out as

$$GA = R \quad (3.2)$$

where G is the following diagonal matrix:

$$G = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & d_1 & 0 & \cdots & 0 \\ 0 & 0 & d & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & d \end{bmatrix}$$

$$d = \text{Var } Q_i - \text{Cov}(Q_i, Q_j) \quad (i, j \neq 0) \quad (3.3)$$

and

$$(v+1)d_1 = v[\text{Var}(Q_0) - 2\text{Cov}(Q_0, Q_i) + \text{Cov}(Q_i, Q_j)] + d \quad (3.4)$$

In such designs $\text{Cov}(Q_0, Q_i)$ is constant independent of i , $\text{Cov}(Q_i, Q_j) = \text{constant}$ and $\text{Var}(Q_i) = \text{constant}$ ($i, j = 1, 2, \dots, v$).

Here, evidently d_1 and d are two eigen values of the variance-covariance matrix of the adjusted totals and d has multiplicity $V - 1$.

The solutions of a_i 's are

$$a_2 = \frac{R_2}{d_1}$$

and

$$a_i = \frac{R_i}{d} \quad (i \neq 1, 2)$$

Accordingly, $\text{Var}(a_2) = \frac{\sigma^2}{d_1}$,

and $\text{Var}(a_i) = \frac{\sigma^2}{d}$

and the covariance between the estimates of any two a_i 's ($i = 2, \dots, V$) is zero ensuring additivity of information recovered on them. For finding efficiency we take completely randomised design with the same replications of the treatments as in the efficiency balanced design.

If $a_2 = l_{20} t_0 + l_{21} t_1 + \dots + l_{2v} t_v$ then for a completely randomised design the estimate of a_2 is

$$\hat{a}_2 = l_{20} \frac{T_0}{r_0} + l_{21} \frac{T_1}{r_1} + \dots + l_{2v} \frac{T_v}{r_v}$$

where T_0, T_1 , etc., stand for the treatment totals.

$$\text{Var}(\hat{a}_2) = \sigma^2 \left(\frac{l_{20}^2}{r_0} + \frac{l_{21}^2}{r_1} + \dots + \frac{l_{2v}^2}{r_v} \right).$$

The information recovered relative to CR design is thus

$$d_1 \left(\frac{l_{20}^2}{r_0} + \frac{l_{21}^2}{r_1} + \dots + \frac{l_{2v}^2}{r_v} \right)$$

Evidently this ratio has to be the same for every contrast as the designs are efficiency balanced. This is actually the efficiency of the designs and is denoted by E . If we take a_2 as $1/\sqrt{2}(t_1 - t_2)$ its efficiency is d/r where r is the replication of any of the treatments t_1 to t_v .

As in these designs the estimates of mutually orthogonal contrasts of t 's are also mutually orthogonal, the total loss of information is $(V - 1)(1 - E)$.

We have expressed below the total loss of information as a function of parameters of different series of efficiency balanced designs. There is one series of designs in which there is no extra treatment. For this series V has to be replaced by $(V - 1)$ in relations 3.1 to 3.4 of pp. 67-77 of Das and Ghosh [1].

3.1 Efficiency Balanced Designs with $V + 1$ Treatments

Efficiency balanced designs are obtained by reinforcing a BIB design with parameters V, b, r, k, λ with one extra treatment and a number of n extra blocks $n > 0$. The extra treatment does not occur in any of the original blocks but occurs $(r - \lambda)/\lambda$ times in each of the extra blocks, and each of the other treatment occurs once in each of the extra blocks. The two values of the eigen values are :

$$d_1 = \frac{(V + 1) n (r - \lambda)}{rk} \quad \text{and} \quad d = rE + n.$$

Where E is the efficiency factor of the original B.I.B. design and $r + n$ is the replication of each of the treatments, t_1 to t_v .

The efficiencies of this series of design = $d/(r + n)$

$$= \frac{rE + n}{r + n}$$

Total loss of information = $V(1 - (rE + n)/(r + n))$

$$= Vr \left(\frac{1 - E}{r + n} \right)$$

3.2 Efficiency Balanced Designs with V Treatments

Efficiency balanced designs are obtained by reinforcing a B.I.B. design with parameters v, b, r, k and λ by taking any number of n extra blocks ($n \geq 0$). Any particular treatment, say, first one occurs r/λ times in each of the extra blocks' and each of other treatment occurs once in each of the extra blocks. The two eigen values are

$$d_1 = \frac{v}{k} (n + \lambda) \quad \text{and} \quad d = \frac{v}{k} + n$$

and efficiency of this series of design = $d/(r + n)$

$$= \frac{nk + v\lambda}{K(r + n)}$$

Total loss of information is thus $(V - 1)(1 - (nk + v\lambda)/(nk + kr))$

$$= \frac{(V - 1)(r - \lambda)}{K(r + n)}$$

Illustrations. These findings are illustrated below by taking two typical designs.

1. We take the design with parameters $v = 4, b = 6, r = 3, k = 2, \lambda = 1$ and $n = 1$. Its incidence matrix P and the orthogonal matrix C are shown below :

		Treatments				
		0	1	2	3	4
$P' =$	[0	1	1	0	0
		0	1	0	1	0
		0	1	0	0	1
		0	0	1	1	0
		0	0	1	0	1
		0	0	0	1	1
		2	1	1	1	1
]					
$C' =$	[$\frac{1}{\sqrt{5}}$	$\frac{1}{\sqrt{5}}$	$\frac{1}{\sqrt{5}}$	$\frac{1}{\sqrt{5}}$	$\frac{1}{\sqrt{5}}$
		$\frac{4}{\sqrt{20}}$	$\frac{-1}{\sqrt{20}}$	$\frac{-1}{\sqrt{20}}$	$\frac{-1}{\sqrt{20}}$	$\frac{-1}{\sqrt{20}}$
		0	$\frac{1}{\sqrt{4}}$	$\frac{1}{\sqrt{4}}$	$\frac{-1}{\sqrt{4}}$	$\frac{-1}{\sqrt{4}}$
		0	$\frac{1}{\sqrt{4}}$	$\frac{-1}{\sqrt{4}}$	$\frac{1}{\sqrt{4}}$	$\frac{-1}{\sqrt{4}}$
		0	$\frac{1}{\sqrt{4}}$	$\frac{-1}{\sqrt{4}}$	$\frac{-1}{\sqrt{4}}$	$\frac{1}{\sqrt{4}}$
]					

For this design $C'(N - P'Q^{-1}P)C$ is the following diagonal matrix

$$G = \frac{1}{6} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 18 & 0 & 0 \\ 0 & 0 & 0 & 18 & 0 \\ 0 & 0 & 0 & 0 & 18 \end{bmatrix}$$

Thus,

$$d_1 = \frac{10}{6}, d = \frac{18}{6} = 3$$

and efficiency $E = 3/4$

Total loss of information = $V(1 - E) = 4 \times 1/4 = 1$.

Illustration 2. We have taken below an efficiency balanced design where $n_{ij} \leq 1$. The parameters, incidence matrix and the C matrix for the design are shown below :

$$P = \begin{array}{c} \text{Treatments} \\ \hline \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \\ \hline \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \end{array} \quad C' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} \\ \frac{1}{\sqrt{20}} & \frac{-1}{\sqrt{20}} & \frac{-1}{\sqrt{20}} & \frac{-1}{\sqrt{20}} & \frac{-1}{\sqrt{20}} \\ 0 & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{-1}{\sqrt{4}} & \frac{-1}{\sqrt{4}} \\ 0 & \frac{1}{\sqrt{4}} & \frac{-1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{-1}{\sqrt{4}} \\ 0 & \frac{1}{\sqrt{4}} & \frac{-1}{\sqrt{4}} & \frac{-1}{\sqrt{4}} & \frac{1}{\sqrt{4}} \end{bmatrix}$$

Here, $v = 4$, $b = 6$, $r = 3$, $k = 2$, $n = 2$, $p = 1$, $q = 1$ and $\lambda = 1$

$$G = \frac{1}{6} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 30 & 0 & 0 & 0 \\ 0 & 0 & 26 & 0 & 0 \\ 0 & 0 & 0 & 26 & 0 \\ 0 & 0 & 0 & 0 & 26 \end{bmatrix}$$

Thus

$$d_1 = \frac{30}{6} \quad \text{and} \quad d = \frac{26}{6}$$

Efficiency, $E = 13/15$; Total loss of information = $V(1 - E)$

$$= 4 \left(1 - \frac{13}{15} \right) = \frac{8}{15}$$

As the efficiency has been obtained relative to a completely randomised design, the total loss obtained here need not agree with the result of Tyagi. But if we find the efficiency relative to a randomised block design

with replication equal to the average of replications, that is, with $\bar{r} = 26/5$ the efficiency for the estimate of the contrast a_2 is d_1/\bar{r} . The efficiency for the contrasts a_3 , a_4 and a_5 is d/\bar{r} each.

Hence total loss of information is

$$\left(1 - \frac{d_1}{\bar{r}}\right) + 3\left(1 - \frac{d}{\bar{r}}\right)$$

$$= \frac{14}{26} \text{ after substituting values of } d_1 \text{ and } d \text{ and } \bar{r} \text{ and}$$

$$\text{on simplification.}$$

REFERENCES

- [1] Das, M. N. and Ghosh, D. K. (1985): Balancing incomplete block designs, *Sankhya : The Indian Journal of Statistics*, 1985, 47 (1): Series B, 67-77.
- [2] Das, M. N. (1958): Reinforced incomplete block designs, *J. Indian Soc. Agricultural Statistics*, 10 : 73-77.
- [3] Dey, A. and Singh, M. (1980): Some series of efficiency balanced designs, *Austra. J. Statist.*, 22 (3) : 364-367.
- [4] Kshirsagar, A. M. (1957): A note on the total relative loss of information in any design. *Calc. Stat. Ass. Bull.*, 7 : 78-81.
- [5] Puri, P.D. and Nigam, A. K. (1975a): On patterns of efficiency balanced designs. *J. Roy. Stat. Soc., B*, 37 : 357-58.
- [6] Tyagi, B. N. (1969): A note on the relative loss of information in confounded designs, *Journal Indian Society of Agri. Statistics*, 59-61.
- [7] Yates, F. (1937): The design and analysis of factorial experiments, *Imp. Bur. Soil. Sci. Technical Concern.* No. 35.