Sankhyā : The Indian Journal of Statistics
1981, Volume 43, Series B, Pt. 2, pp. 222-227.

ROBUSTNESS OF THREE DIMENSIONAL DESIGNS, I

By SUBIR GHOSH
Indian Statistical Institute

SUMMARY. In this paper we study a robustness property of a latin square design (LSD) of order s in the sense that, when any t observations are missing the design remains connected w.r.t. treatment, row, and column factors. We show that a LSD of order s ($\geqslant 4$) is robust against missing of any $s-1$ observations.

1. Introduction

The robustness of designs against incomplete data was considered in papers of Ghosh (1978, 1979a, 1979b). In this paper we consider the robustness of a three dimensional design, LSD of order s which eliminates heterogeneity in two directions. Various optimum properties of LSD are well known, see Kiefer (1958) and Wald (1943). This paper gives a further property of robustness of such design. We show in a LSD of order s ($\geqslant 4$), when any $s-1$ observations are missing the resulting design remains connected w.r.t. 3 factors treatment, row, and column. We observe that LSD's of order 2 and 3 are not robust in the above sense. We shall consider the robustness of other 3 dimensional designs in subsequent communications.

2. Robustness of LSD

Consider a 3 dimensional design with u rows, w columns, and v treatments. Suppose y_{ijk} is the observation corresponding to the k-th treatment in the i-th row and j-th column, τ_k is the effect of the k-th treatment, r_i is the effect of the i-th row, c_j is the effect of the j-th column, and μ is the general mean. Then the model is

$$y_{ijk} = \mu+r_i+c_j+\tau_k+e_{ijk},$$
$$i = 1, \ldots, u,\ j = 1, \ldots, w,\ k = 1, \ldots, v, \qquad \ldots \quad (1)$$

where the experimental error e_{ijk}s' are uncorrelated random variables with mean 0 and variance σ^2.

Definition 1 : A linear function of treatment effects $\sum_{k=1}^{v} C_k\tau_k$, C_k ($k = 1, \ldots, v$) are real numbers, is said to be a contrast if $\sum_{k=1}^{v} C_k = 0$.

Definition 2 : Two treatments k and l are said to be connected if $\tau_k - \tau_l$ is estimable under the model (1).

Definition 3 : A 3 dimensional design is said to be connected w.r.t. treatment factor if every pair of treatments is connected.

Similar definitions can be given for row and column connectedness of a 3 dimensional design. We now recall a definition of Bose and Srivastava (1964), Srivastava and Anderson (1970).

Definition 4 : A 3 dimensional design is said to be completely connected if it is connected w..r.t. 3 factors treatment, row, and column.

If some observations in a completely connected 3 dimensional design is missing then the resulting design may or may not be connected w.r.t. treatment factor or completely connected.

Definition 5 : A completely connected 3 dimensional design is said to be robust against missing of any t (a positive integer) observations if the design obtained by omitting any t observations remains connected w.r.t. treatment factor Clearly, $t \leqslant \min(\rho_1, \ldots, \rho_v) - 1$, where the treatment k is replicated ρ_k times.

Definition 6 : A completely connected 3 dimensional design is said to be strongly robust against missing of any t observations if the design obtained by omitting any t observations remains completely connected.

A LSD of order s, LSD(s), is a completely connected 3-dimensional design. Consider the following latin squares of order 2 and 3.

1	2
2	1

1	2	3
2	3	1
3	1	2

Suppose the observations corresponding to treatment 1 in cell (1, 1) of LSD(2) and treatments 1 and 3 in cells (1, 1) and (2, 2) of LSD(3) are missing. Then $\tau_1 - \tau_2$ is not estimable in the resulting design from LSD(2) and only contrast $\tau_1 + \tau_3 - 2\tau_2$ is estimable in the resulting design from LSD(3). Thus a LSD(s), $s = 2, 3$, is not robust against missing of any $s-1$ observations.

Consider a LSD(s), $s > 4$, with $(k_1, k_2, \ldots, k_s)$ and $(l_1, l_2, \ldots, l_s)$ as two rows. Let $(l_{w_1}, l_{w_2}, \ldots, l_{w_s}) = (k_1, k_2, \ldots, k_s)$, where $(w_1, w_2, \ldots, w_s)$ is a permutation of $(1, 2, \ldots, s)$ with the cycle structure $1^{a_1}, 2^{a_2}, \ldots, s^{a_s}$, $a_1 = 0$, a_i $(i = 2, \ldots, s)$ are nonnegative integers satisfying $a_1+2a_2+\ldots+sa_s = s$.

Theorem 1 : *If the rows $(k_1, \ldots, k_s)$ and $(l_1, \ldots, l_s)$ are such that $a_2 \neq 0$ and $a_j = 0$ $(j \neq 2)$, then there are at least a_2 linearly independent contrasts of τs' which are estimable.*

Proof: Suppose $k_u = l_{u+1}$, $k_{u+1} = l_u$, $u = 1, 3, \ldots, s-1$. Then clearly $\tau_{k_u} - \tau_{k_{u+1}}$, $u = 1, 3, \ldots, s-1$, are estimable. Therefore we have $a_2 = (s/2)$ independent contrasts of τ's which are estimable. This completes the proof.

Theorem 2 : *If the rows $(k_1, \ldots, k_s)$ and $(l_1, \ldots, l_s)$ are such that at least one a_j $(j \neq 2)$ is nonzero, then there are at least $s-a_2-1$ linearly independent contrasts of τs'.*

Proof: Suppose $(w_1, w_2, \ldots, w_p)$ is a permutation of $(1, 2, \ldots, p)$ and the permutation is a p-cycle. The elements in the two rows corresponding to u-th and w_u-th columns are (k_u, k_{w_u}) and (l_u, l_{w_u}), $u = 1, \ldots, p$. We know $l_{w_u} = k_u$. It follows from the model (1) that the contrasts $2\tau_{k_u} - \tau_{l_u} - \tau_{k_{w_u}}$, $u = 1, \ldots, p$, are estimable. Notice that

$$\sum_{u=1}^{p} \{2\tau_{k_u} - \tau_{l_u} - \tau_{k_{w_u}}\} = 0.$$

It can be checked that among these p contrasts $p-1$ are linearly independent. Therefore from different cycles we get $a_3+2a_3+\ldots+(s-1)a_s$ independent estimable contrasts.

Suppose $(w_1, \ldots, w_p)$ and $(w_{p+1}, \ldots, w_{p+q})$ are permutations of $(1, \ldots, p)$ and $(p+1, \ldots, p+q)$ and the permutations are p (> 2) and q (> 2) cycles. The elements in the two rows corresponding to u-th and $p+v$-th columns, $1 \leqslant u \leqslant p$, $1 \leqslant v \leqslant q$, are (k_u, k_{p+v}) and (l_u, l_{p+v}). The contrast $\tau_{k_u} + \tau_{l_{p+v}} - \tau_{l_u} - \tau_{k_{p+v}}$ is estimable and is independent of the earlier contrasts. Thus a_i i-cycles $(i > 2)$ give a_i-1 independent contrasts, a_j j-cycles $(j > 2)$

give a_j-1 independent contrasts, an i-cycle and a j-cycle give a contrast as illustrated earlier. Therefore we get a set of $a_3+a_4+\ldots+a_s-1$ independent estimable contrasts which are independent of the earlier contrasts. Hence the total number of independent estimable contrasts is

$$[a_2+2a_3+\ldots+(s-1)a_s]+[a_3+\ldots+a_s-1],$$

i.e.,

$$s-a_2-1.$$

This completes the proof.

Suppose a set of $s-1$ observations are missing in a LSD(s). It is clear that there are at least one row and one column in which no observation is missing. Consider one such row, say $(k_1, k_2, \ldots, k_s)$. There must be another row in which at most one observation is missing. Denote it by $(l_1, l_2, \ldots, l_s)$.

Case 1 : Suppose no observation is missing in $(l_1, \ldots, l_s)$.

(i) Let at least one $a_j \neq 0$ $(j \neq 2)$. Considering two rows $(k_1, \ldots, k_s)$ and $l_1, \ldots, l_s)$, it follows from Theorem 2 that there are $s-a_2-1$ independent estimable contrasts. Corresponding to a_2 2-cycles, we get from all columns a_2 sets containing 2 columns each and another set containing $s-2a_2$ other columns. Thus we divide all columns into a_2+1 sets of columns. Consider $s-2$ rows other than $(k_1, \ldots, k_s)$ and $(l_1, \ldots, l_s)$. In any two of a_2+1 sets there is a row in which at least one observation is not missing in each set. We take one treatment from each set, the observations for which are not missing, and the treatments which belong to the same columns as those in the row $(k_1, \ldots, k_s)$. From these 4 treatments we find a contrast which is estimable. It can be checked that this contrast is independent of $s-a_2-1$ earlier contrasts. The sets $(i, i+1)$, $i = 1, \ldots, a_2$, will give a_2 independent estimable contrasts which are independent of $s-a_2-1$ contrasts. Thus we have a set of $s-1$ independent estimable contrasts.

(ii) Let $a_j = 0$ $(j \neq 2)$. Considering the two rows $(k_1, \ldots, k_s)$ and $(l_1, \ldots, l_s)$ it follows from Theorem 1 that there are a_2 independent contrasts. Considering a_2 2-cycles and arguing similar to above we find a_2-1 independent estimable contrasts which are independent of a_2 contrasts stated earlier. We thus have $2a_2-1$, i.e., $s-1$ independent estimable contrasts.

Case 2 : Consider the situation where there is one observation missing in $(l_1, \ldots, l_s)$ and there is no other block besides $(k_1, \ldots, k_s)$ in which no observation is missing.

Theorem 3 : *If one observation in $(l_1, \ldots, l_s)$ is missing and no observation in $(k_1, \ldots, k_s)$ is missing and furthermore $a_2 \neq 0$, then*

(a) *there are at least $s-a_2-1$ independent estimable contrasts of τs' in case at least one $a_j \neq 0$ $(j \neq 2)$,*

(b) *there are at least a_2 independent estimable contrasts in case $a_j = 0$ $(j \neq 2)$.*

Proof : Similar to Theorems 1 and 2.

Theorem 4 : *If one observation in $(l_1, \ldots, l_s)$ is missing and no observation in $(k_1, \ldots, k_s)$ is missing and, moreover, $a_2 = 0$, then there are at least $s-2$ independent contrasts of τs'.*

Proof : Similar to Theorem 2.

If the design and the pattern of $s-1$ missing observations be so that $a_2 \neq 0$ then arguing similarly as in missing of no observation in $(l_1 \ldots l_s)$ the existence of $s-1$ independent estimable contrasts can be shown. Thus we need to show the existence of $s-1$ independent estimable contrasts in case $a_2 = 0$. In Theorem 4 we have already shown the existence of $s-2$ independent estimable contrasts. We therefore need to show the existence of a single estimable contrast which is independent of other $s-2$ contrasts. Suppose the observation corresponding to l_1 is missing. Clearly the observation corresponding to at least one of the treatments k_1 and l_1 is not missing in all other $s-2$ rows. There must be two rows and two columns so that in their intersecting cells either (l_1, α) and (β, l_1), or, (k_1, γ) and (δ, k_1) are occurring. Thus $2\tau_{l_1}-\tau_\alpha-\tau_\beta$ or $2\tau_{k_1}-\tau_\gamma-\tau_\delta$ is estimable. It can be checked that any of $2\tau_{l_1}-\tau_\alpha-\tau_\beta$ and $2\tau_{k_1}-\tau_\gamma-\tau_\delta$ is independent of the other $s-2$ contrasts. Thus we obtain a set of $s-1$ independent estimable contrasts.

Therefore we state the following theorem :

Theorem 5 : *A LSD of order s $(\geqslant 4)$ is robust against missing of any $s-1$ observations.*

Corollary 1 : *A LSD of order s $(\geqslant 4)$ is strongly robust against missing of any $s-1$ observations.*

Proof : When any $s-1$ observations are missing, there are at least one observation not missing in each row and each column. The proof is now clear.

References

Bose, R. C. and Srivastava, J. N. (1964) : Multidimensional partially balanced designs and their analysis, with applications to partially balanced factorial fractions. *Sankhyā*, Ser A, **26**, 145-368.

Ghosh, S. (1978) : On robustness of optimal balanced resolution V plans (submitted for publication).

——— (1979a) : On robustness of designs against incomplete data. *Sankhyā* Ser B, Pts. 3 and 4, **40**, 204-208.

——— (1979b) : A robustness property of BIBD (Submitted for publication).

Kiefer, J. C. (1958) : On the nonrandomized optimally and randomized nonoptimality of symmetrical designs. *Ann. Math. Stat.*, **29**, 675-699.

Srivastava, J. N. and Anderson, D. A. (1970) : Some basic properties of multidimensional partially balanced designs. *Ann. Math. Stat.*, **41**, 1438-1445.

Wald, A. (1943) : On the efficient design of statistical investigations. *Ann. Math. Stat.*, **14**, 134-140.

Paper received : July, 1979.

Revised : May, 1980.