*Jour. Ind. Soc. Ag. Statistics*
*56(2), 2003 : 107-116*

# Robustness of Youden Square Designs Against Missing Data

Krishan Lal, V.K. Gupta and Lalmohan Bhar
*Indian Agricultural Statistics Research Institute, New Delhi-110012*
(Received : May, 1998)

SUMMARY

This article investigates the robustness of Youden square and Latin square designs against missing data. The robustness has been investigated against the loss of any t(≥1) observations in a column/row and for the loss of any two observations in the design as per connectedness criterion. The robustness has also been investigated as per the A-efficiency criterion.

*Key words* : A-efficiency, Connectedness, Latin square designs, Robustness, Row-column designs, Youden square designs.

## *1. Introduction*

The loss of data in field experiments is a common phenomenon, which at times leads to some serious problems. Firstly, the residual design may become disconnected and it may not be possible to make all possible paired comparisons through the design. Secondly, although the residual design may be connected, the loss of efficiency may be so high that even an optimal design may render to be less efficient. A design that is capable of absorbing such shocks after the loss of data is called a robust design. An experimenter will prefer a design that remains connected even after the loss of any t(≥1) observations (hereafter called as Criterion 1 robustness studied by Ghosh [5]). The experimenter may further prefer that the A-efficiency of the residual design, as compared to the original design, is not too small (hereafter called as Criterion 2 robustness studied by John [11], Gupta and Srivastava [8] and Dey [2]).

The robustness of incomplete block designs against missing data has been investigated in the literature from different angles (see *e.g.* Hedayat and John [10], John [11], Ghosh ([5], [6], [7]), Baksalary and Tabis [1], Dey and Dhall [4], Srivastava *et al.* ([16], [17], [18]), Gupta and Srivastava [8], Mukerjee and Kageyama [14], Dey [2], Dey *et al.* [3] and Lal *et al.* [13]). The robustness of Latin square designs (LSD) against missing data has been studied by Ghosh [6] as per Criterion 1. It has been shown that a Latin square design of size $v \geq 4$ is robust against the loss of any v-1 observations. Srivastava *et al.* ([16], [17]) observed that Youden square designs (YSD), obtainable through symmetrical balanced incomplete block (BIB) designs listed in Raghavarao [15], are almost robust against the loss of all observations in one column as per Criteria 1 and 2.

This article gives further results on robustness of Youden square designs, as per Criteria 1 and 2 against the loss of any $t(\geq 1)$ observations in a column/row and for the loss of any two observations in the design pertaining to same or different treatments. YSDs obtained from symmetric BIB designs and LSDs have been identified for robustness against the loss of any $t(\geq 1)$ observations in a column/row and any two observations in the design. Srivastava *et al.* ([16], [17]) studied particular case when t = p, where p is column size.

Throughout the present investigation we shall deal with real matrices and vectors. Denote an n-component vector of all unities by $\mathbf{1}_n$, an identity matrix of order n by $\mathbf{I}_n$ and $m \times n$ matrix of all ones by $\mathbf{J}_{m\times n}$, $\mathbf{J}_{m\times m}$ is simply denoted by $\mathbf{J}_m$. Further, $\mathbf{A}'$, $\mathbf{A}^-$ and $\mathbf{A}^+$ will respectively denote the transpose, a generalized inverse (g-inverse) and the Moore-Penrose inverse of a matrix $\mathbf{A}$.

## 2. *Condition for Robustness*

### 2.1 Some Preliminaries

We consider the usual homoscedastic, additive, fixed effects linear model under two blocking systems as

$$\mathbf{M} \equiv \{\mathbf{y}, \mathbf{X}\,\theta\,, \sigma^2 \mathbf{I}_n\} \tag{2.1}$$

where $\mathbf{X} = [\mathbf{1}_n \quad \Delta' \quad \mathbf{D}_1' \quad \mathbf{D}_2']$, $\theta = [\mu \quad \tau' \quad \psi' \quad \chi']'$, $\tau = (\tau_1, \ldots, \tau_v)'$ is a $v \times 1$ vector of treatment effects, $\psi = (\psi_1, \ldots, \psi_p)'$ is a $p \times 1$ vector of row effects, $\chi = (\chi_1, \ldots, \chi_q)'$ is a $q \times 1$ vector of column effects, $\mathbf{y}$ is $n \times 1$ vector of observations, $\mu$ is the general mean, $\Delta'$ is $n \times v$ design matrix of observations vs treatment effects, $\mathbf{D}_1'$ is $n \times p$ design matrix of observations vs row effects, $\mathbf{D}_2'$ is $n \times q$ design matrix of observations vs column effects and $\mathbf{e} = (e_1, \ldots, e_n)'$ is an $n \times 1$ vector of non-observable random variables, each distributed with mean zero and constant variance $\sigma^2$. Our interest is in estimating the parameter vector $\tau$. Partition $\theta$ as $\theta = (\theta_1' \vdots \theta_2')'$, where $\theta_1 = \tau$ and $\theta_2 = [\mu \quad \psi' \quad \chi']'$ and $\mathbf{X} = [\mathbf{X}_1 \vdots \mathbf{X}_2]$ with $\mathbf{X}_1$ and $\mathbf{X}_2$ are $n \times v$ and $n \times (p+q+1)$ matrices, respectively.

Also $\Delta' \mathbf{1}_v = \mathbf{D}_1' \mathbf{1}_p = \mathbf{D}_2' \mathbf{1}_q = \mathbf{1}_n$. Let $\mathbf{r} = \Delta \mathbf{1}_n = (r_1, \ldots, r_v)'$, $\mathbf{k}_1 = \mathbf{D}_1 \mathbf{1}_n = (k_{11}, \ldots, k_{1p})'$ and $\mathbf{k}_2 = \mathbf{D}_2 \mathbf{1}_n = (k_{21}, \ldots, k_{2q})'$ denote respectively, the vector of replications, the vector of row sizes and the vector of column sizes and let $\mathbf{R} = \Delta\Delta' = \operatorname{diag}(r_1, \ldots, r_v)$, $\mathbf{K}_1 = \mathbf{D}_1\mathbf{D}_1' = \operatorname{diag}(k_{11}, \ldots, k_{1p})$ and $\mathbf{K}_2 = \mathbf{D}_2\mathbf{D}_2' = (k_{21}, \ldots, k_{2q})$. Also let $\mathbf{N}_1 = \Delta\mathbf{D}_1'$ be the $v \times p$ treatments vs rows incidence matrix, $\mathbf{N}_2 = \Delta\mathbf{D}_2'$ be the $v \times q$ treatments vs columns incidence matrix and $\mathbf{W} = \mathbf{D}_1\mathbf{D}_2'$ be the $p \times q$ rows vs columns incidence matrix. We then have $\mathbf{W1}_q = \mathbf{k}_1 = \mathbf{N}_1'\mathbf{1}_v$, $\mathbf{W}'\mathbf{1}_p = \mathbf{k}_2 = \mathbf{N}_2'\mathbf{1}_v$, $\mathbf{N}_1\mathbf{1}_p = \mathbf{r} = \mathbf{N}_2\mathbf{1}_q$ and $\mathbf{r}'\mathbf{1}_v = \mathbf{k}_1'\mathbf{1}_p = \mathbf{k}_2'\mathbf{1}_q = n$. The information matrix for obtaining the best linear unbiased estimators (BLUE) of the estimable parametric functions of $\tau$, after eliminating the other nuisance parameters, is

$$\mathbf{C}_\tau = \Delta\mathbf{P}_{\mathbf{x}_2}\Delta' \tag{2.2}$$

where $\mathbf{P}_{\mathbf{x}_2} = \mathbf{I} - \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'$

$$= \mathbf{I} - \mathbf{D}_1'\mathbf{K}_1^{-1}\mathbf{D}_1 - (\mathbf{D}_1'\mathbf{K}_1^{-1}\mathbf{W} - \mathbf{D}_2')\mathbf{F}^-(\mathbf{W}'\mathbf{K}_1^{-1}\mathbf{D}_1 - \mathbf{D}_2)$$

$$\mathbf{F} = \mathbf{K}_2 - \mathbf{W}'\mathbf{K}_1^{-1}\mathbf{W}$$

It is easy to verify that $\mathbf{P}_{\mathbf{x}_2}\mathbf{1} = \mathbf{0}$

We assume that the design d considered here is connected and Rank $(\mathbf{C}_\tau) = v - 1$. Suppose now that any $t(\geq 1)$ observations are lost in d. Let $d_t$ be the residual design and we consider the model of t missing observations as

$$\mathbf{M}_t \equiv \{\mathbf{Ay}, \mathbf{AX}\theta, \sigma^2\mathbf{A}\} \tag{2.3}$$

where $\mathbf{A} = \mathbf{I} - \mathbf{U}$, $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]$ and $\mathbf{u}_j = (0, 0, \ldots, 1(j^{th}), 0, 0, \ldots, 0)'$ is an n-component vector with a 1 in the $j^{th}$ position if $j^{th}$ observation is lost and 0 elsewhere, $j = 1, \ldots, n$. Under this model usual **C**-matrix simplifies to

$$\mathbf{C}_{\tau(t)} = \Delta\mathbf{A}\Delta' - \Delta\mathbf{AX}_2(\mathbf{X}_2'\mathbf{AX}_2)^-\mathbf{X}_2'\mathbf{A}\Delta \tag{2.4}$$

We consider now another model in which we devote an extra parameter to each missing observation as

$$\mathbf{M}_z \equiv \{\mathbf{y}, \mathbf{Z}\delta, \sigma^2\mathbf{I}_n\} \tag{2.5}$$

where $$\mathbf{Z} = [\mathbf{X}\ \ \mathbf{U}],\ \delta = [\theta'\ \ \gamma']' \tag{2.6}$$

It can be easily verified that the **C**-matrix derived from model $\mathbf{M}_z$ is identical to the matrix given in (2.4) (see, Lal *et al.* [13]). Henceforth, we shall use the model $\mathbf{M}_z$ for studying the robustness. The following relationship between $\mathbf{C}_\tau$ and $\mathbf{C}_{\tau(t)}$ is known (Lal *et al.* [13]).

*Lemma 2.1.* $\mathbf{C}_\tau = \mathbf{C}_{\tau(t)} + \mathbf{VC}_*^-\mathbf{V}'$, where $\mathbf{C}_* = \mathbf{UP}_{\mathbf{x}_2}\mathbf{U}$ is non-negative definite and $\mathbf{V} = \Delta\mathbf{P}_{\mathbf{x}_2}\mathbf{U}$.

2.2 Conditions for Robustness

In this section we present the condition for robustness of two-way heterogeneity design d under the model (2.1) against missing data. For robustness against the loss of any $t(\geq 1)$ observations as per Criterion 1, the following result is due to Lal *et al.* [13].

*Lemma 2.2.* The design d is Criterion 1 robust against the loss of any t observations if the smallest eigenvalue of $\mathbf{C}_\tau$ is strictly greater than the largest eigenvalue of $\mathbf{VC}_*^-\mathbf{V}'$.

Once it is known that a design is Criterion 1 robust it is of interest to examine the efficiency of the residual design relative to original design and decide robustness on the basis of Criterion 2. The A-efficiency of residual design with respect to original design is given by

$$E = \frac{\text{Sum of reciprocal of non-zero eigen values of } \mathbf{C}_\tau}{\text{Sum of reciprocal of non-zero eigen values of } \mathbf{C}_{\tau(t)}} = \frac{\text{trace}[\mathbf{C}_\tau^+]}{\text{trace}[\mathbf{C}_{\tau(t)}^+]} \tag{2.7}$$

Now using Theorem 2 of Dey [2], we have

$$\mathbf{C}_{\tau(t)}^+ = \mathbf{C}_\tau^+ + \mathbf{C}_\tau^+\mathbf{VC}_*^{-1/2}[\mathbf{I}_n - \mathbf{C}_*^{-1/2}\mathbf{V}'\mathbf{C}_\tau^+\mathbf{VC}_*^{-1/2}]^{-1}\mathbf{C}_*^{-1/2}\mathbf{V}'\mathbf{C}_\tau^+ \tag{2.8}$$

so $\quad \text{trace}[\mathbf{C}_{\tau(t)}^+] = \text{trace}[\mathbf{C}_\tau^+] + g(\alpha)$ (2.9)

where $\quad g(\alpha) = \text{trace}[\mathbf{C}_\tau^+\mathbf{VC}_*^{-1/2}(\mathbf{I}_n - \mathbf{C}_*^{-1/2}\mathbf{V}'\mathbf{C}_\tau^+\mathbf{VC}_*^{-1/2})^{-1}\mathbf{C}_*^{-1/2}\mathbf{V}'\mathbf{C}_\tau^+]$

$$= \text{trace}[(\mathbf{I}_n - \mathbf{C}_*^{-1/2}\mathbf{V}'\mathbf{C}_\tau^+\mathbf{VC}_*^{-1/2})^{-1}\mathbf{C}_*^{-1/2}\mathbf{V}'\mathbf{C}_\tau^+\mathbf{C}_\tau^+\mathbf{VC}_*^{-1/2}] \tag{2.10}$$

Thus $$E = \frac{\text{trace}[\mathbf{C}_\tau^+]}{\text{trace}[\mathbf{C}_\tau^+] + g(\alpha)} = \left[1 + \frac{g(\alpha)}{\text{trace}[\mathbf{C}_\tau^+]}\right]^{-1} \tag{2.11}$$

Let the design be variance balanced and $\eta$ is the unique non-zero eigenvalue of $\mathbf{C}_\tau$ with multiplicity v – 1, then $\mathbf{C}_\tau^+ = \frac{1}{\eta}\mathbf{I}_v$. (2.12)

For computing efficiency, we need to calculate the eigenvalues of $\mathbf{C}_*^{-1/2}\mathbf{V}'\mathbf{VC}_*^{-1/2}$. Since the non-zero eigen values of $\mathbf{C}_*^{-1/2}\mathbf{V}'\mathbf{VC}_*^{-1/2}$, $\mathbf{VC}_*^-\mathbf{V}'$ and $\mathbf{C}_*^-\mathbf{V}'\mathbf{V}$ are same, we work out the eigenvalues of $\mathbf{VC}_*^-\mathbf{V}'$ or $\mathbf{C}_*^-\mathbf{V}'\mathbf{V}$ instead of $\mathbf{C}_*^{-1/2}\mathbf{V}'\mathbf{VC}_*^{-1/2}$.

Let $e_1 \geq e_2 \geq \ldots \geq e_m$ are the m($\leq$t) positive eigenvalues of $\mathbf{VC}_*^-\mathbf{V}'$ or $\mathbf{C}_*^{-1}\mathbf{V}'\mathbf{V}$, where t($\geq$1) is the number of observations lost, then for a variance balanced design we have

$$E = \left[1 + \frac{1}{v-1}\sum_{i=1}^{m}\frac{e_i}{(\eta - e_i)}\right]^{-1} \tag{2.13}$$

## *3. Applications*

In Section 2 we have given some conditions for robustness of a design d for a general two-way heterogeneity setting in which all the classifications are possibly non-orthogonal. Although these conditions are applicable to any connected design for a two-way heterogeneity setting, further simplifications are possible to identify robust designs if we restrict our attention to specific designs.

### 3.1 Youden Square Designs

A row-column design is a two-way heterogeneity design with rows vs columns classification orthogonal i.e. $\mathbf{W} = \frac{\mathbf{k}_1\mathbf{k}_2'}{n}$. In the class of row-column designs the simplest one are Latin square designs with v treatments, v rows and v columns (p = q = v) and Youden square designs with p rows, q (= v) columns and v treatments. In Youden design $p \times q$ array has one observation in each cell i.e. $W = J_{pq}$ and $\mathbf{P}_{x_2} = \mathbf{I} - \frac{\mathbf{D}_1'\mathbf{D}_1}{q} - \frac{\mathbf{D}_2'\mathbf{D}_2}{p} + \frac{\mathbf{J}_n}{n}$, $n = pq$

Without any loss of generality, we suppose that the n observations are arranged in such a way that the first p observations come from first column, the next p observations come from second column and so on, and the last p observations come from the $q^{th}$ column. Then we have $\mathbf{D}_1' = \mathbf{1}_q \otimes \mathbf{I}_p$, $\mathbf{D}_2' = \mathbf{I}_q \otimes \mathbf{1}_p$ and $\mathbf{P}_{x_2} = \phi_p \otimes \phi_q = \phi_q \otimes \phi_p$ (if we change the order of observations of rows and column), where $\otimes$ stands for Kronecker product of two matrices and $\phi_p = \mathbf{I} - \frac{1}{p}\mathbf{J}_p$ and $\phi_q = \mathbf{I} - \frac{1}{q}\mathbf{J}_q$.

Similarly we partition $\Delta'$ as $\Delta' = [\Delta_1 \quad \Delta_2 \ \dots \ \Delta_q]'$, where $\Delta_j'$ is a $p \times v$ $(0-1)$ matrix for treatments; j = 1, 2, …, q. Thus **C**-matrix can be written as

$$\mathbf{C}_\tau = \Delta\mathbf{P}_{x_2}\Delta' = \Delta(\phi_q \otimes \phi_p)\Delta' \tag{3.1}$$

A Youden square design can always be constructed from a symmetrical BIB design. Also Latin square is a particular case of Youden square design when p = v. The non-zero eigenvalues of **C**-matrix of a Youden square design and a Latin square design are $\frac{\rho v(p-1)}{v-1}$ and $\rho v$, respectively with multiplicity $v - 1$, where $\rho$ is the number of times the design is replicated. Now we study the robustness of Youden square and Latin square designs when missing observations follow specific pattern as per both the criteria of robustness.

3.2. Robustness of Youden Square Designs

We consider the following cases of missing observations.

*Case 1.* When any $t(1 \le t \le p)$ observations are lost in a column

Suppose that $t(1 \le t \le p)$ observations belonging to one of the columns of design d are lost. Without loss of generality, let the missing observations pertain to first t treatments in the first column of d. Then it is easy to see that

$$\mathbf{C}_*^- \mathbf{V'V} = \begin{bmatrix} \frac{\rho v}{\rho v - 1} I_t & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \text{ if } 1 \le t \le p - 1 \tag{3.2}$$

$$= \begin{bmatrix} \frac{\rho v}{\rho v - 1} \phi_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \text{ if } t = p \tag{3.3}$$

The non-zero eigenvalues of $\mathbf{C}_*^- \mathbf{V'V}$ are $\frac{\rho v}{\rho v - 1}$ with multiplicity t, if $1 \le t \le p - 1$ and $\frac{\rho v}{\rho v - 1}$ with multiplicity $p - 1$, if $t = p$. Then by the use of Lemma 2.2 we arrive at

*Theorem 3.1.* Youden square designs are Criterion 1 robust against the loss of any $t(1 \le t \le p)$ observations in a column if $p > \frac{(\rho + 1)v - 2}{(\rho v - 1)}$.

*Corollary 3.1.* All the 62 Youden square designs obtained from symmetric BIB designs listed in Raghavarao [15], Kageyama [12] and Hall [9] are Criterion 1 robust for the loss of any t observations in a column, for all values of $\rho$.

The A-efficiency of the residual design from (2.13) is simplified as

$$E(t) = \left[1 + \frac{t}{(p-1)(\rho v - 1) - (v - 1)}\right]^{-1}, \qquad \text{if } 1 \le t \le p - 1 \tag{3.4}$$

$$E(p) = \left[1 + \frac{p-1}{(p-1)(\rho v - 1) - (v - 1)}\right]^{-1}, \qquad \text{if } t = p \tag{3.5}$$

*Remark 3.1.* The efficiency of residual design in comparison to original design decreases as $t(1 \le t \le p)$, the number of missing observations increases in a column in Youden square designs. This efficiency is minimum when $t = p$ or $p - 1$.

*Remark 3.2.* By symmetry, the robustness of a residual design when any t observations are lost in a row is same as when any t observations are lost in a column.

*Case 2.* When any two observations are missing in a Youden square design

When any two observations are missing in a Youden square design, the situation can be treated by separating into the following patterns

*Case 2(a).* When any two missing observations belong to the same column

This is a particular case of Case 1 when t = 2 and is represented by $E_1(2)$.

*Case 2(b).* When both the missing observations belong to the same treatment in different columns/rows

Suppose without any loss of generality, that the two missing observations pertain to the first position in the first column and the second position in the second column and both the observations pertain to the same treatment, then we have

$$\mathbf{C}_* = \begin{bmatrix} \frac{(p-1)(\rho v-1)}{\rho vp} & \mathbf{0}' & \frac{1}{\rho vp} & \mathbf{0}' \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{1}{\rho vp} & \mathbf{0}' & \frac{(p-1)(\rho v-1)}{\rho vp} & \mathbf{0}' \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{3.6}$$

On simplifying we get the two non-zero eigenvalues of $\mathbf{C}_*^{-}\mathbf{V}'\mathbf{V}$ as $e_1 = \frac{\rho vp}{a+1}(b+c)$ and $e_2 = \frac{\rho vp}{a-1}(b-c)$, where $a = (p-1)(\rho v-1)$, $b = \frac{(p-1)}{p}$, $c = \left(\frac{p-1}{p}\right)^2 + \frac{\lambda-1}{p^2}$ and $\lambda$ is the number of columns in which each pair of treatments occurs together in a symmetric BIB design from which Youden square is obtained. With these eigenvalues we compute the A-efficiency $E_2(2)$ from (2.13).

*Case 2(c).* When both the observations belong to different treatments in different columns/rows

Without any loss of generality, let the two missing observations belong to the first position in the first column and the second position in the second column and both observations pertain to different treatments. In the present case, no unique eigenvalue of $\mathbf{C}_*^{-}\mathbf{V}'\mathbf{V}$ can be obtained because with interchange of columns of Youden square design, the non-zero eigenvalues of $\mathbf{C}_*^{-}\mathbf{V}'\mathbf{V}$ also change. However a lower bound of efficiency can be obtained if we are able to work out the maximum eigenvalues of $\mathbf{C}_*^{-}\mathbf{V}'\mathbf{V}$. For this we work out the eigenvalues of $\mathbf{C}_*^{-}\mathbf{V}'\mathbf{V}$ for all possible structures of $\mathbf{C}_*^{-}\mathbf{V}'\mathbf{V}$. Interestingly, it is found that the eigenvalues of $\mathbf{C}_*^{-}\mathbf{V}'\mathbf{V}$ are maximum when both the observations belong to the same treatment. Thus the efficiency calculated for the Case 2(b), [$E_2(2)$] provides the lower bound for the present case.

In Table 1 the values of efficiencies for different cases of missing observations in Youden square designs obtained from symmetric BIB designs listed in Raghavarao [15], Kageyama [12] and Hall [9] for $\rho = 1$and $\rho = 2$ are presented in Table 1. It reveals that for the loss of one observation in a column, the loss of efficiency is more than 20% in design R2 and for the loss of a complete column the efficiencies are greater than 80% except for four designs, viz. R2, R4, R8 and R10, when $\rho = 1$. Also loss of efficiency increases when the number of missing observations increases in a column and it is maximum when $t = p$ or $p - 1$. Youden square designs are fairly robust against the loss of all observations in a column. For the loss of two observations, except for the designs R2, R4 and R10 the efficiencies are greater than 80%, when $\rho = 1$. The efficiencies are always greater than 80% for all the cases studied here, when $\rho = 2$. Thus we conclude that Youden square designs are fairly robust against the loss of any 2 observations in the design as per Criterion 2.

**Table 1.** A-efficiency of the residual design when the original design is a Youden square design

| | E | ≥ 0.99 | 0.95 - 0.99 | 0.90 - 0.95 | 0.80-0.90 | < 0.80 |
|---|---|---|---|---|---|---|
| $\rho = 1$ | E(1) | 44 | 13 | 2 | 2 | 1 |
| | E(p) | 12 | 29 | 8 | 9 | 4 |
| | $E_1(2)$ | 39 | 14 | 4 | 3 | 2 |
| | $E_2(2)$ | 38 | 14 | 4 | 3 | 3 |
| $\rho = 2$ | E(1) | 52 | 9 | 1 | 0 | 0 |
| | E(p) | 26 | 26 | 8 | 2 | 0 |
| | $E_1(2)$ | 44 | 14 | 3 | 1 | 0 |
| | $E_2(2)$ | 44 | 13 | 3 | 2 | 0 |

### 3.3 Latin Square Designs

Latin square design is a particular case of a Youden square design when $p = v$. Thus the expressions for calculating the efficiencies are same as those of Youden square designs with the change $p = v$. However explicit expression can be obtained for the case when the two missing observations pertain to different treatments in different rows or columns. Without loss of generality we assume that the first missing observation belongs to the first treatment at the first position of first column and the second observation to any other treatment at the second position in the second column. Thus the matrix $\mathbf{C}_*$ takes the following form

$$= \begin{bmatrix} \frac{(v-1)(\rho v-1)}{\rho v^2} & \mathbf{0}' & \frac{1}{\rho v^2} & \mathbf{0}' \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{1}{\rho v^2} & \mathbf{0}' & \frac{(v-1)(\rho v-1)}{\rho v^2} & \mathbf{0}' \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

On simplifying, we get the two non-zero eigen-values of $\mathbf{C}_*^-\mathbf{V}'\mathbf{V}$ as

$$e_1 = \frac{\rho v^2(v-2)[\rho(v-1)-1]}{(\rho v-1)^2(v-1)^2-1}$$

and $$e_2 = \frac{\rho v^2(\rho v^2-\rho v-v+2)}{(\rho v-1)^2(v-1)^2-1}$$

Thus the A-efficiency of the residual design is given by

$$E_3(2) = \left[1+\frac{1}{v-1}\sum_{i=1}^{2}\frac{e_i}{(\rho v-e_i)}\right]^{-1}$$

**Table 2**. A-efficiency of the residual design when the original design is a Latin square design

| | ρ = 1 | | | | | ρ = 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| v | E(1) | E(p) | $E_1(2)$ | $E_2(2)$ | $E_3(2)$ | E(1) | E(p) | $E_1(2)$ | $E_2(2)$ | $E_3(2)$ |
| 3 | 0.67 | 0.50 | 0.50 | 0.33 | * | 0.89 | 0.80 | 0.80 | 0.78 | 0.77 |
| 4 | 0.86 | 0.67 | 0.75 | 0.67 | 0.71 | 0.95 | 0.86 | 0.90 | 0.89 | 0.90 |
| 5 | 0.92 | 0.75 | 0.86 | 0.82 | 0.85 | 0.97 | 0.89 | 0.94 | 0.94 | 0.94 |
| 6 | 0.95 | 0.80 | 0.91 | 0.89 | 0.91 | 0.98 | 0.91 | 0.96 | 0.96 | 0.96 |
| 7 | 0.97 | 0.83 | 0.94 | 0.93 | 0.94 | 0.99 | 0.92 | 0.97 | 0.97 | 0.97 |
| 8 | 0.98 | 0.86 | 0.96 | 0.95 | 0.95 | 0.99 | 0.93 | 0.98 | 0.98 | 0.98 |
| 9 | 0.98 | 0.88 | 0.97 | 0.96 | 0.97 | 0.99 | 0.94 | 0.99 | 0.98 | 0.99 |
| ≥10 | ≥0.99 | ≥0.89 | ≥0.97 | ≥0.97 | ≥0.97 | ≥0.99 | ≥0.95 | ≥0.99 | ≥0.99 | ≥0.99 |

Note: * represents that the design is not Criterion 1 robust.

Similar to the Youden square designs, the A-efficiencies for Latin square designs are presented in Table 2. It is evident from Table 2 that the loss of efficiency is very high for Latin square design for v = 3 in all the cases. The loss of efficiency is also high when v = 4 for the loss of any two observations or all the observations in a column. The efficiency is quiet high for all the cases for $v \geq 5$. Also the efficiency decreases as the number of missing observations increases in a row or column and it is minimum when t = v or v – 1. The efficiencies are always greater than 80%, except for the loss of any two observations or a complete column/row for v = 3, when ρ = 2. Thus Latin square designs for v ≥ 5 are fairly robust against the loss of any two observations and loss of any number of observations in a column.

ACKNOWLEDGEMENT

The authors are thankful to the referees for the useful suggestions that have helped in improvement in the presentation of results.

REFERENCES

[1] Baksalary, J.K. and Tabis, Z. (1987). Conditions for robustness of block designs against unavailability of data. *J. Statist. Plann. Inf.*, **16**, 49-54.

[2] Dey, A. (1993). Robustness of block designs against missing data. *Statistica Sinica*, **3**, 219-231.

[3] Dey, A., Midha, K. Chand and Buchthal, D.C. (1996). Efficiency residual design under the loss of observations in a block design. *Jour. Ind. Soc. Agril. Stat.*, **49**, 237-248.

[4] Dey, A. and Dhall, S.P. (1988). Robustness of augmented BIB designs. *Sankhya*, **B50**, 376-381.

[5] Ghosh, S. (1979). On robustness of designs against incomplete data. *Sankhya*, **B40**, 204-208.

[6] Ghosh, S. (1981). Robustness of three dimensional designs. *Sankhya*, **B43**, 222-227.

[7] Ghosh, S. (1982). Robustness of BIB designs against the unavailability of data. *J. Statist. Plann. Inf.*, **6**, 29-32.

[8] Gupta, V. K. and Srivastava, R. (1992). Investigations on robustness of block designs against missing observations. *Sankhya*, **B54**, 100-105.

[9] Hall, M. (1986). *Combinatorial Theory*. Wiley Interscience, New York.

[10] Hedayat, A. and John, P.W.M. (1974). Resistant and susceptible BIB designs. *Ann. Statist.*, **2**, 148-158.

[11] John, P.W.M. (1976). Robustness of incomplete block designs. *Ann. Statist.*, **4**, 960-962.

[12] Kageyama, S. (1972). A survey of resolvable solution of balanced incomplete block designs. *Int. Statist. Rev.*, **40(3)**, 269-273.

[13] Lal, K., Gupta, V. K. and Bhar, L. (2001). Robustness of designed experiments against missing data. *J. Appl. Statist.*, **28**, 63-79.

[14] Mukerjee, R. and Kageyama, S. (1990). Robustness of group divisible designs. *Commu. Statist. –Theory and Methods*, **19**, 3189-3203.

[15] Raghavarao, D. (1971). *Construction and Combinatorial Problems in Design of Experiments*. Wiley, New York.

[16] Srivastava, R., Gupta, V. K. and Dey, A. (1990). Robustness of some designs against missing observations. *Commu. Statist. –Theory and Methods*, **19**, 121-126.

[17] Srivastava , R. , Gupta, V. K. and Dey, A. (1991). Robustness of some designs against missing data. *J. Appl. Statist.*, **18**, 313-318.

[18] Srivastava, R., Prasad, R. and Gupta, V. K. (1996). Robustness of block designs for making test treatments - control comparisons against a missing observation. *Sankhya*, **B58**, 407-413.