

A New Class of Efficiency Balanced Two-Way Elimination of Heterogeneity Designs Useful in Sericulture

Sanpei Kageyama, Satyabrata Pal¹, Anurup Majumder¹ and Arunava Ghosh²
(Received September 10, 2003)

In this paper a new class of efficiency balanced two-way elimination of heterogeneity designs, derived from F-squares, has been obtained. These designs have potential applications in sericulture.

Keywords: F-squares, Generalized row-column design, orthogonal generalized row-column design

1. Introduction

Two-way elimination of heterogeneity designs are useful in field experiments if the fertility gradient exists in two directions. The available literatures on the constructions of such designs are, indeed, plenty. The most common useful two-way elimination of heterogeneity design is the latin square design. A generalization of a latin square design is the F-square design, due to Federer. In this paper, the definition of an F-derivative design is introduced. It has been shown that such designs can be used in sericultural experiments wherein better strains of silkworms are screened from a number of competitive strains. This paper begins with an enunciation of a real life problem encountered in sericultural research. The motivation of the paper stems from an effort to have a solution to the sericulture-research problem.

Farmers engaged in sericulture are interested in such strains of silkworm which can produce cocoon in the least possible time after the formation of egg owing to the economic advantage that they can have for being able to process the cocoon for extraction of silk. The procedure of selection of better strains from a set of v strains to be compared is governed by observing the above time period and subsequently the judgement is done on the basis of the criterion of the above-mentioned least possible time. Experimental researches based on this objective are conducted in sericultural farms. Perennial Mulberry trees are planted in orchards, usually, in squared layouts, the field having two-way heterogeneity along rows and columns, and the leaves of such plants are used for feeding the larvae of silkworms. Normally, owing to the requirement of such experiments, latin squares/F-squares (of type $F(m, \alpha)$) for allowing replications within rows and columns are chosen. Thus, m^2 Mulberry trees, planted under a two-way heterogeneity set-up in the orchard in conformity with the existing two-dimensional heterogeneity, simulating a frequency square set-up with v different above-mentioned strains, to be compared, following an $F(m, \alpha)$ set-up, each strain being replicated α times in each row and in each column.

In sericulture, the silkworm larvae need be reared with utmost care since they are very susceptible to disease. Rearing medium (house) need be thoroughly cleaned and disinfected. In the above rearing medium, m^2 boxes, each with the same aged 20,000 eggs, are taken for

¹Bidhan Chandra Krishi Viswavidyalaya, Mohanpur 741 252, India

²Uttar Banga Krishi Viswavidyalaya, Pundibari 736 165, India

rearing. Each of the m^2 boxes is designated by a level (i, j) , the identity of its placement, $i, j = 1, 2, \dots, m$. These boxes kept in rearing houses are exposed to a totally controlled environment except for the presence of unidentifiable/uncontrollable factors operating in each box. Eggs of v different strains (pure/monohybrid/polyhybrid) of *Bombyx mori* are taken for comparison of the length of their larval periods (from egg to cocoon formation) and are placed in the above-mentioned m^2 boxes allowing α replications in each row and in each column in respect of each strain, generating a lay-out of an $F(m, \alpha)$ design. Leaves from each of m^2 Mulberry plants are then supplied to the eggs in the corresponding box out of the m^2 boxes considered above. Thus, the positional effects (rows and columns) of Mulberry trees exist implicitly in the corresponding box. The usual analysis of F-squares is valid on the observations (lengths of larval period) recorded from the boxes.

The motivation of the present communication is delineated below. It is, indeed, possible that v strains can be compared in m^2 boxes under an $F(m, \alpha)$ design arranged in m rows and m columns. However, in the orchard, $m - x$ (not m rows, $x \neq 0$) and m columns, $m(m - x)$ plots, may be available in many practical situations. In such cases, the larvae contained in the first set of x boxes in the first column may be fed with the leaves of the tree positioned in the first row and first column in the orchard. Technically, it means that the larvae in the boxes having the serial positions, $(1, 1), (2, 1), \dots, (x, 1)$, respectively, are fed with the leaves of the tree occurring in the first $(1, 1)$ position, first row and first column, in the orchard. Similarly, the larvae in the boxes having the serial positions, $(1, j), (2, j), \dots, (x, j)$, respectively, are fed with the leaves of the tree occurring in the position $(1, j), j = 2, 3, \dots, m$, in the orchard. Thus, the feeding system adopted here tentamounts to a new arrangement of the m rows in an $F(m, \alpha)$ design such that the first x rows of the $F(m, \alpha)$ design are adjoined to form a single row in which x treatments are placed in each of the m cells of the newly-constructed row. The arrangement in the rest $m - x$ rows remains unaltered. Obviously, it is understood that the $m - x$ row effects and m column effects do also influence the response variable, the length of the larval period, as before. The x observations occurring in each of the m cells in the first row remain independent owing to the fact that the x boxes are different. In fact, the constructed design mentioned above is seen to be an orthogonal generalized row-column design. A variant of the above design is also obtained which is found to be an efficiency balanced orthogonal generalized row-column design.

2. Preliminaries and definitions

The following preliminaries and definitions are presented for the sake of clarity. In what follows, only connected two-way elimination of heterogeneity designs are considered. Most of the symbols have been recalled from Pal (1977) and Dean and Voss (1999).

Throughout this paper, $\mathbf{1}_v$ is a $v \times 1$ column vector with ones, \mathbf{I}_v is the identity matrix of order v , $\mathbf{J}_{s \times t} = \mathbf{1}_s \mathbf{1}_t'$ and especially $\mathbf{J}_v = \mathbf{1}_v \mathbf{1}_v'$.

Definition 2.1. A generalized row-column (GRC) design, with v treatments, b rows and b' columns, is defined as a two-way elimination of heterogeneity design in which the estimates of row effects ignoring treatment effects are orthogonal to the estimates of column effects ignoring treatment effects (see also Chakrabarti, 1962).

It is known (Pal, 1977) that a sufficient condition for a two-way elimination of heterogeneity design to be a GRC design is $\mathbf{N}^* = n^{-1} \mathbf{k} \mathbf{q}'$, where \mathbf{N}^* is the row-column incidence matrix, n is the total number of observations, $\mathbf{k} = [k_1, k_2, \dots, k_b]'$ (row-size vector) and $\mathbf{q} = [q_1, q_2, \dots, q_{b'}]'$ (column-size vector). In case a GRC design has $\mathbf{N}^* = \mathbf{J}_{b \times b'}$, a $b \times b'$

row-column (RC) design is obtained. For a GRC design, the coefficient matrix, C , under the usual two-way elimination of heterogeneity model can be written as $C\alpha = Q$, where $C = r^\delta - Nk^{-\delta}N' - \underline{N}q^{-\delta}\underline{N}' + n^{-1}rr'$, $Q = Nk^{-\delta}B - \underline{N}q^{-\delta}\underline{C}' + n^{-1}Gr$, α is the estimate of the vector of treatment effects, N^* , N and \underline{N} are the incidence matrices of row vs column effects, treatment vs row effects, and treatment vs column effects, respectively, $r = [r_1, r_2, \dots, r_v]'$ (replication vector), $r^\delta = \text{diag}\{r_1, r_2, \dots, r_v\}$, $k^{-\delta} = \text{diag}\{1/k_1, 1/k_2, \dots, 1/k_b\}$, $q^{-\delta} = \text{diag}\{1/q_1, 1/q_2, \dots, 1/q_{b'}\}$, Q is the vector of adjusted treatment totals, and T , B and \underline{C} are the treatment-total vector, row-total vector and column-total vector, respectively, and G is the grand total.

An important matrix M_0 (see Caliński and Kageyama, 2000, 2002) is defined below.

$$\begin{aligned} M_0 &= r^{-\delta}Nk^{-\delta}N' + r^{-\delta}\underline{N}q^{-\delta}\underline{N}' - 2n^{-1}\mathbf{1}_v r' \\ &= (r^{-\delta}Nk^{-\delta}N' - n^{-1}\mathbf{1}_v r') + (r^{-\delta}\underline{N}q^{-\delta}\underline{N}' - n^{-1}\mathbf{1}_v r') \\ &= M_{01} + M_{02} \text{ (say)}. \end{aligned}$$

Definition 2.2. A GRC design or an RC design is said to be orthogonal if the estimates of the column effects γ after eliminating treatment and row effects are orthogonal to the estimates of row effects β after eliminating treatment and column effects (see also Pal, 1977).

It is mentioned (cf. Pal, 1977) that $\text{Cov}(\beta, \gamma) = 0$ (orthogonality condition) holds under the satisfaction of the condition $N^* = \underline{N}'r^{-\delta}N$.

Definition 2.3. A GRC design or an RC design is said to be efficiency balanced (EB) if the relation $M_0s = \mu s$, where s is any vector in a set of $v - 1$ independent vectors of treatment effect contrasts and μ is the non-zero unique eigenvalue of the information matrix M_0 .

In this case, $M_0 = \mu(I_v - n^{-1}\mathbf{1}_v r')$ holds.

Definition 2.4. Let $A = (a_{ij})$ be an $n \times n$ matrix and $\Sigma = \{c_1, c_2, \dots, c_m\}$ be an ordered set of distinct elements of A . In addition, suppose that for each k ($= 1, 2, \dots, m$), c_k appears precisely λ_k (≥ 1) times in each row and in each column of A . Then A is called a frequency square or simply, an F-square on Σ of order n with frequency $(\lambda_1, \lambda_2, \dots, \lambda_m)$.

A matrix A is said to be an $F(n; \lambda_1, \lambda_2, \dots, \lambda_m)$ square array if A is an F-square of order n with frequency $(\lambda_1, \lambda_2, \dots, \lambda_m)$, $\sum_j \lambda_j = n$. An F-square array produces an F-square design. An $F(n; \lambda^m)$ square represents an $F(n; \lambda, \lambda, \dots, \lambda)$ square array, while an $F(n; \lambda_1^2, \lambda_2^3, \lambda_3)$ represents an $F(n; \lambda_1, \lambda_1, \lambda_2, \lambda_2, \lambda_2, \lambda_3)$. In particular, in an $F(n; \lambda^m)$ square, m is determined uniquely by n and λ , i.e., $n = m\lambda$. Hence such a square is represented by $F(n; \lambda)$.

Example 2.1. Let $\Sigma = \{1, 2, 3\}$. Then the following array is an $F(6; 2^3)$ on Σ .

$$\begin{bmatrix} 1 & 2 & 3 & 3 & 2 & 1 \\ 2 & 3 & 1 & 1 & 3 & 2 \\ 3 & 1 & 2 & 2 & 1 & 3 \\ 3 & 1 & 2 & 2 & 1 & 3 \\ 2 & 3 & 1 & 1 & 3 & 2 \\ 1 & 2 & 3 & 3 & 2 & 2 \end{bmatrix}.$$

The above array is also an F-square design with the three treatments 1, 2, 3, each replicated

twice in each row and in each column.

The definitions of variance balanced and efficiency balanced block/RC designs can also be found in Pal (1977) and Caliński and Kageyama (2000, 2002).

3. Construction of efficiency balanced orthogonal generalized row-column designs

In what follows, two methods of construction, one on orthogonal GRC designs and the other on efficiency balanced orthogonal GRC designs, are presented. These designs are constructed from F-square designs.

Method 3.1. Let us take any successive $p (\geq 2)$ rows/columns, say, j_1, j_2, \dots, j_p , of an F-square design of order m and these are combined into a single row/column. The set of p observations at m column/row positions constitute in all m sets of p values corresponding to m column/row positions forming a new design (obtained after merging) with its first row/column having p observations at each of the m column/row positions in that row/column. The other rows/columns of the basic design remain unchanged. The new design so obtained contains $m - p + 1$ rows/columns and m columns/rows. The process of combining can be extended for a second and third, etc., sets of rows/columns.

The procedure of construction explained in Method 3.1 can be well understood through the following. Let us take an F-square design $F(6; 2^3)$ on $\Sigma = \{1, 2, 3\}$, as in Example 2.1. Now if we merge the first two rows, then the following design is obtained.

$$\begin{bmatrix} (1, 2) & (2, 3) & (3, 1) & (3, 1) & (2, 3) & (1, 2) \\ 3 & 1 & 2 & 2 & 1 & 3 \\ 3 & 1 & 2 & 2 & 1 & 3 \\ 2 & 3 & 1 & 1 & 3 & 2 \\ 1 & 2 & 3 & 3 & 2 & 2 \end{bmatrix}.$$

The new design has 5 rows and 6 columns, each of the treatments in the first row has 2 treatments instead of one treatment.

By following the above method, an orthogonal GRC design can be developed in which any treatment effect contrast is estimated with full information.

Proof (by construction)

Case I: Let A be an $F(m, 2)$ square design with v treatments, each replicated twice in each row and in each column, so $m = 2v$. The first p rows are now merged into only one row with p treatments in each cell (at each column position) of the first row. Recalling the notations, the new design has the following (structural) vectors and matrices:

$$\mathbf{r} = 4v\mathbf{1}_v, \mathbf{k} = [2pv, 2v\mathbf{1}'_{m-p}]', \mathbf{q} = 2v\mathbf{1}_m, \mathbf{N} = [2p\mathbf{1}_v : 2\mathbf{1}_v\mathbf{1}'_{m-p}],$$

$$\mathbf{N}^* = \begin{bmatrix} p\mathbf{1}'_{2v} \\ \mathbf{1}_{m-p}\mathbf{1}'_{2v} \end{bmatrix}.$$

Let $\underline{\mathbf{N}}$ be the $v \times 2v$ treatment-column incidence matrix. Thus, $\underline{\mathbf{N}} = 2\mathbf{1}_v\mathbf{1}'_{2v}$. Then it follows that

$$\mathbf{N}'\mathbf{r}^{-\delta}\underline{\mathbf{N}} = \begin{bmatrix} p\mathbf{1}'_{2v} \\ \mathbf{1}_{m-p}\mathbf{1}'_{2v} \end{bmatrix} \quad \text{and} \quad \frac{1}{n}\mathbf{k}\mathbf{q}' = \mathbf{N}'\mathbf{r}^{-\delta}\underline{\mathbf{N}} \quad \text{with} \quad n = m^2.$$

According to Definition 1.1, the above-modified F-square is an orthogonal GRC design.

Case II: Let B be an F-square design as $F(m; \lambda_1^u, \lambda_2^{v-u})$ with v treatments for m and u being positive integers and $m > v$. Here, $u\lambda_1 + (v-u)\lambda_2 = m$ and the total number of observations is $n = m^2$.

Let $\mathbf{k} = m\mathbf{1}_m$, row-size vector, $\mathbf{q} = m\mathbf{1}_m$, column-size vector, and further $\mathbf{r} = [m\lambda_1\mathbf{1}'_u, m\lambda_2\mathbf{1}'_{v-u}]'$, replication vector. Further let us impose Method 3.1 to the F-square design B to construct an orthogonal GRC design. The corresponding (structural) vectors and matrices of the modified B (after imposition) are given as $\mathbf{k} = [pm, m\mathbf{1}'_{m-p}]'$, row-size vector, $\mathbf{q} = m\mathbf{1}_m$, column-size vector, and $\mathbf{r} = [m\lambda_1\mathbf{1}'_u, m\lambda_2\mathbf{1}'_{v-u}]'$, replication vector. The treatment-row incidence matrix is given by

$$\mathbf{N} = \begin{bmatrix} p\lambda_1\mathbf{1}_u & \lambda_1\mathbf{1}_u\mathbf{1}'_{m-p} \\ p\lambda_2\mathbf{1}_{v-u} & \lambda_2\mathbf{1}_{v-u}\mathbf{1}'_{m-p} \end{bmatrix},$$

while the treatment-column incidence matrix is given by

$$\underline{\mathbf{N}} = \begin{bmatrix} \lambda_1\mathbf{1}_u\mathbf{1}'_m \\ \lambda_2\mathbf{1}_{v-u}\mathbf{1}'_m \end{bmatrix}.$$

Then it follows that

$$\mathbf{N}'\mathbf{r}^{-\delta}\underline{\mathbf{N}} = \begin{bmatrix} \frac{p(u\lambda_1+(v-u)\lambda_2)}{u\lambda_1+(v-u)\lambda_2} \mathbf{1}'_m \\ \frac{p(u\lambda_1+(v-u)\lambda_2)}{m} \mathbf{1}_{m-p}\mathbf{1}'_m \end{bmatrix} = \begin{bmatrix} p\mathbf{1}'_m \\ \mathbf{1}_{m-p}\mathbf{1}'_m \end{bmatrix}.$$

Again, it holds that

$$\frac{1}{n}\mathbf{k}\mathbf{q}' = \begin{bmatrix} p\mathbf{1}'_m \\ \mathbf{1}_{m-p}\mathbf{1}'_m \end{bmatrix} = \mathbf{N}'\mathbf{r}^{-\delta}\underline{\mathbf{N}}.$$

Thus the modified design B is an orthogonal GRC design. \square

Method 3.2. Consider a design \mathcal{D} which is an $F(m, \alpha)$, with $m = v\alpha$. If a particular row/column is deleted from such a design, the residual design \mathcal{D}^* so formed is called an F-derivative RC design. The deleted row/column makes the column/row incomplete. However, the deleted row/column does not disturb the equality of the number of paired occurrences of any two treatments (the proof is given below). The design \mathcal{D}^* has the following properties:

Without loss of generality, let the last row of the above $F(m, \alpha)$ design be deleted. Then the treatment-row two-way design of \mathcal{D}^* is complete as each row contains all treatments, each treatment occurring α times and hence an orthogonal design, while the treatment-column two-way design of \mathcal{D}^* is a variance balanced (two-way) design, as shown below:

If $\underline{\mathbf{N}}^*$ is the treatment-column incidence matrix of \mathcal{D}^* , then $\underline{\mathbf{N}}^* = (\alpha - 1)\mathbf{1}_v\mathbf{1}'_m + \mathbf{K}^*$, where \mathbf{K}^* is the $v \times m$ incidence matrix of α identical replicates of a balanced incomplete block design with parameters $v, b = v, r = v - 1, k = v - 1, \lambda = v - 2$. In fact,

$$\begin{aligned} \underline{\mathbf{N}}^*(\underline{\mathbf{N}}^*)' &= [(\alpha - 1)\mathbf{1}_v\mathbf{1}'_m + \mathbf{K}^*][(\alpha - 1)\mathbf{1}_v\mathbf{1}'_m + \mathbf{K}^*]' \\ &= m(\alpha - 1)^2\mathbf{J}_v + (\alpha - 1)\mathbf{1}_v\mathbf{1}'_m(\mathbf{K}^*)' + (\alpha - 1)\mathbf{K}^*\mathbf{1}_m\mathbf{1}'_v + \mathbf{K}^*(\mathbf{K}^*)' \\ &= \alpha v(\alpha - 1)^2\mathbf{J}_v + \alpha(\alpha - 1)(v - 1)\mathbf{J}_v + \alpha(\alpha - 1)(v - 1)\mathbf{J}_v \\ &\quad + \alpha\mathbf{I}_v + \alpha(v - 2)\mathbf{J}_v \\ &= \alpha\mathbf{I}_v + \alpha^2(m - 2)\mathbf{J}_v. \end{aligned}$$

Thus, the two-way treatment-column design is variance balanced and the constancy ($= \alpha^2(m - 2)$) of the occurrences of the pair of treatments is maintained in the said two-way design.

If N^* and N^{**} are the incidence matrices between treatment vs row and row vs column, respectively, of \mathcal{D}^* , then it is shown that $N^{**} = (N^*)'r^{-\delta}\underline{N}$, where $N^{**} = \mathbf{1}_{m-1}\mathbf{1}'_m$, $r^{-\delta} = [\alpha(m-1)]^{-1}I_v$ and $N^* = \alpha\mathbf{1}_v\mathbf{1}'_{m-1}$.

Now, $(N^*)'r^{-\delta}\underline{N} = [\alpha(m-1)]^{-1}\alpha\mathbf{1}_{m-1}\mathbf{1}'_v\{(\alpha-1)\mathbf{1}_v\mathbf{1}'_m + K^*\} = [v(\alpha-1)/(m-1)]\mathbf{1}_{m-1}\mathbf{1}'_m + [(v-1)/(m-1)]\mathbf{1}_{m-1}\mathbf{1}'_m = \mathbf{1}_{m-1}\mathbf{1}'_m$.

The coefficient matrix C^* of \mathcal{D}^* can be written as $C^* = \alpha(m-1)I_v - J_v - (m-1)^{-1}\{\alpha I_v + \alpha^2(m-2)J_v\} + \{\alpha^2(m-1)^2/[m(m-1)]\}J_v = \alpha I_v + \alpha^2(m-2)J_v$. Thus the F-derivative RC design is a variance balanced orthogonal RC design.

Lastly, Method 3.1 is employed on the first p rows of the design \mathcal{D}^* , the design so obtained is denoted by \mathcal{D}^{**} . Then \mathcal{D}^{**} has the following (structural) vectors and matrices. The replication, row-size and column-size vectors, treatment-row, treatment-column, row-column incidence matrices, are written as follows: $r = r\mathbf{1}_v$, $k = \alpha v[p, \mathbf{1}'_{h-1}]'$, $q = (m-1)\mathbf{1}_m$, $r = \alpha(m-1)$, $m = \alpha v$, $h = m - p$, $n = m(m-1)$, n_{ij} being the frequency of occurrences of the i th treatment in the j th column, $i = 1, 2, \dots, v$; $j = 1, 2, \dots, m$, $\sum_i n_{ij} = m-1$ for all j ; $\sum_j n_{ij} = r$ for all i ; $\sum_j n_{ij}n_{i'j} = \Lambda = \alpha^2(m-2)$ for all i and $i'(\neq i) = 1, 2, \dots, v$, $\sum_j n_{ij}^2 = (m-2)\alpha^2 + \alpha = R$ for all i , $N = [p\alpha\mathbf{1}_v : \alpha\mathbf{1}_v\mathbf{1}'_{m-p-1}]$ and $\underline{N} = (n_{ij})$ of size $v \times m$. Then it follows that

$$N'r^{-\delta}\underline{N} = \frac{\alpha(m-1)}{r} \begin{bmatrix} p\mathbf{1}'_m \\ \mathbf{1}_{h-1}\mathbf{1}'_m \end{bmatrix} = \begin{bmatrix} p\mathbf{1}'_m \\ \mathbf{1}_{h-1}\mathbf{1}'_m \end{bmatrix} = \frac{1}{n}kq'.$$

Thus \mathcal{D}^{**} is an orthogonal GRC design. Further recall that $M_0 = M_{01} + M_{02}$. Hence, for the design \mathcal{D}^{**} , it follows that

$$\begin{aligned} M_0 &= (r^{-\delta}Nk^{-\delta}N' - n^{-1}\mathbf{1}_vr') + (r^{-\delta}\underline{N}q^{-\delta}\underline{N}' - n^{-1}\mathbf{1}_vr') \\ &= [\alpha(m-1)^2]^{-1}\{(R-\Lambda)I_v + \Lambda J_v\} - v^{-1}J_v \\ &= \{\alpha/[\alpha(m-1)^2]\}I_v + \{(m-2)\alpha^2/[\alpha(m-1)^2] - 1/v\}J_v \\ &= (m-1)^{-2}(I_v - v^{-1}J_v). \end{aligned}$$

Thus the \mathcal{D}^{**} gives a new class of efficiency balanced orthogonal GRC design with efficiency $1 - (m-1)^{-2}$. The \mathcal{D}^{**} is also a variance balanced orthogonal GRC design. These designs are useful in the above sericultural experiments and have potential application in clinical trials.

References

- Caliński, T. and Kageyama, S. (2000). *Block Designs: A Randomization Approach, Volume I: Analysis*. Springer, New York.
- Caliński, T. and Kageyama, S. (2002). *Block Designs: A Randomization Approach, Volume II: Design*. Springer, New York.
- Chakrabarti, M. C. (1962). *Mathematics of Design and Analysis of Experiments*. Asia Publishing House, Bombay.
- Dean, A. and Voss, D. (1999). *Design and Analysis of Experiments*. Springer-Verlag, New York.
- Hedayat, A. and Seiden, E. (1970). F-squares and orthogonal F-squares designs: A generalization of Latin square designs. *Ann. Math. Statist.* **41**, 2035-2044.
- Pal, S. (1977). On designs with one-way and two-way elimination of heterogeneity. *Calcutta Statist. Assoc. Bull.* **26**, 79-104.