Indian Journal of Animal Sciences **85** (9): 1046–1050, September 2015/Article

Irregular Sudoku-type designs for animal experiments

ARVIND KUMAR[1], CINI VARGHESE[2], ELDHO VARGHESE[3] and SEEMA JAGGI[4]

Indian Agricultural Statistics Research Institute, PUSA, New Delhi 110 012 India

Received: 15 February 2014; Accepted: 24 March 2015

ABSTRACT

In experiments involving animals, it may be a difficult task for the experimenter to get homogeneous groups of animals. Blocking is commonly practiced to eliminate variability present in experimental material due to some known sources such as age, breed, housing condition, initial body weight, litter size, etc. When there are two non-interacting sources of variation in the experimental material, experimental designs under two-way blocking structure are advised. However, these designs are not appropriate when a third source of variability is present in the experimental material as they may not be able to mark out the third features that tend to clump the animals into a third set of compact groups. Many times the clumps may be irregular, having no uniformity. In this paper, four methods of constructing designs under three way blocking structure for irregular clumps were developed in which two methods give designs with clumps having empty nodes. These designs assure more precise comparison among the treatments under study by removing three known sources of variability present in the experimental material.

Key words: Efficiency factor, Gerechte designs, Irregular clumps, Structurally incomplete, Variance balanced

Statistical designs under two-way blocking structure are advantageously used in case of variation due to two sources among experimental units. However, these designs may not be able to mark out features that tend to clump in compact groups or similar groups that can be formed due to an additional source of heterogeneity in an animal experiment. For example, an experiment is conducted to study the effect of seven feeds on the body weight of a group of animals of different ages put under six different housing conditions. Animals were grouped into 7 different age groups and one each from each age group was allocated to each housing condition. Thus, a total of 42 animals were required for the study. Further, these 42 animals were having different initial body weights. So, 14 clumps were formed based on the initial body weights of the animals. For such situations, experimental designs under three-way blocking structure like the one given below may be advisable.

1 2 4	2 3 5	3 4 6	4 5 7	5 6 1	6 7 2	7 1 3
3 5 6	4 6 7	5 7 1	6 1 2	7 2 3	1 3 4	2 4 5

Sudoku square designs form a popular class of Gerechte designs that are used to separate three sources of variation from the experimental material. Bailey *et al.* (1991, 2008) showed that the regions of a Gerechte design are always formed from adjacent units. In these designs, rows, columns and regions are treated as blocks. Hui-Dong and Ru-Gen (2008) discussed the basic properties of Sudoku squares and a construction procedure. Subramani and Ponnuswamy (2009) gave a method of construction of a class of Sudoku designs. Saba and Sinha (2014) and Sarkar and Sinha (2014) obtained some variants of Sudoku designs.

In designs under three-way blocking structure, sometimes there may be situations where uniform clumps cannot be formed. Four methods of construction of designs under three-way blocking structure for irregular clumps are presented in the next section. First series of designs has irregular clumps as clump size varies. However, the units in a clump are distributed uniformly with respect to the other 2 sources of variability. The remaining 3 series of designs have irregular clumps as the units in the clumps are not distributed uniformly with respect to the other 2 sources of variation.

Further, considering all the 3 sources of variability, some units may not be available for the experiment or they may not be ready for receiving treatment due to some reasons. Designs under three-way blocking structure having empty nodes are suitable for such situations. Out of the 4 methods discussed in this paper, 2 series of designs deals with empty node situations. All the methods result in variance balanced designs. Canonical efficiency factor of the designs obtained in comparison to an orthogonal design having the same number of treatments has also been worked out. Lists of parameters of designs have been listed out along with

Present address: [1]Indian Statistical Service Probationer (arvind.mk702@gmail.com), [2,4]Principal Scientist (cini_v@iasri.res.in, seema.iasri@outlook.com). [3]Scientist (eldho@iasri.res.in).

estimated variance of treatment contrasts and efficiency factor.

MATERIALS AND METHODS

Experimental setup and model

For designs under 3-way blocking structure with v treatments arranged in p rows, q columns and m clumps, the following 4-way classified model with treatments, rows, columns and regions as the 4 classifications, is considered:

$$y_{ijkh} = \mu + t_i + \alpha_j + \beta_k + g_h + e_{ijkh}; \qquad \ldots(1)$$

$$i = 1,2,\ldots,v;\ j = 1,2,\ldots,p;\ k = 1,2,\ldots,q;\ h = 1,2,\ldots,m$$

where, y_{ijkh}, the response from the j^{th} row, k^{th} column and h^{th} region receiving i^{th} treatment; μ, general mean; t_i, i^{th} treatment effect; α_j, j^{th} row effect; β_k, k^{th} column effect; g_h, h^{th} region effect; e_{ijkh}, error term identically and independently distributed and following normal distribution with mean zero and constant variance. Some terms that are used in the subsequent sections are defined below:

Regular/ irregular clumps: If all the clumps belonging to a design under three-way blocking structure are uniform and of same size, i.e., the units in the clumps are distributed uniformly with respect to the other 2 sources of variation and an equal number of units are appearing in each clump, then such designs are said to be regular. Otherwise, they are called irregular. Also, the clumps may be irregular with respect to the incompleteness of experimental units.

Irregular Sudoku-type designs: A design under 3-way blocking structure is said to be irregular Sudoku-type if there exists at least one clump which is irregular.

Structurally incomplete designs: A design under three-way blocking structure is said to be structurally incomplete if there exists at least one unit which does not receive any treatment.

Methods of Construction

General methods of construction of 4 series of designs under 3-way blocking having irregular clumps are described here:

Designs under three-way blocking with irregularly sized clumps obtained using balanced incomplete block designs

Consider the popular balanced incomplete block (BIB) designs for number of treatments v = 4t+3 where v is a prime number and t is any natural number ≥ 1. Obtain an initial block by taking even/ odd powers of primitive element mod v. Obtain another initial block by considering the complementary treatments of this block. As v is a prime number, the complementary block will be of a different size. Juxtapose these initial blocks of one below the other in a vertical manner retaining the identity of the blocks. Develop these initial blocks cyclically mod v in a horizontal direction resulting in 2v blocks. Treating these blocks as regions, a series of designs for three-way blocking structure in v rows, v columns and 2v clumps having v replications can be obtained. v clumps are of size $\frac{(v-1)}{2}$ and another v clumps are of size $\frac{(v+1)}{2}$. For this class of designs the parameters are v (number of treatments) = r (number of replications) = p (number of rows) = q (number of columns), m (number of clumps) = 2v. A general form of information matrix for estimating the elementary contrast pertaining to treatment effects is obtained utilizing the Model 1 as $C_\tau = \frac{v(v-2)}{(v-1)}\left[I_v - \frac{J_v}{v}\right]$ indicating that these designs are found to be variance balanced. The estimated variances of elementary contrasts pertaining to treatment effects (V_{ij}; i,j = 1,2,…,v; i ≠ j) are computed using the SAS code developed and the canonical efficiency factor (E_{can}) of these designs in comparison to an orthogonal design having the same number of treatments assuming the same variance was also computed. Parameters of the designs along with the estimated variance and efficiency factor are listed in Table 1 for v < 25.

Example 1: Let v = 7. Consider the even powers of primitive element to get one initial block as 1, 2, 4. Take the complementary of this block i.e., 3,5,6,7 and place them below the above initial block and develop them cyclically mod 7 to the right. Blocks are taken as clumps and the resultant design has a total number of 14 clumps. This design is variance balanced with efficiency factor 0.8333.

1	2	3	4	5	6	7
2	3	4	5	6	7	1
4	5	6	7	1	2	3
3	4	5	6	7	1	2
5	6	7	1	2	3	4
6	7	1	2	3	4	5
7	1	2	3	4	5	6

For this design parameters are v = r = p = q = 7, m = 14 and the information matrix is of the form $\mathbf{C}_\tau = 5.8333\mathbf{I}_7 - 0.8333\mathbf{J}_7$

Table 1. List of designs

S. No.	v	r	p	q	m	V_{ij}	E_{can}
1	7	7	7	7	14	0.3428	0.8333
2	11	11	11	11	22	0.2020	0.9000
3	19	19	19	19	38	0.1114	0.9444
4	23	23	23	23	46	0.0911	0.9545

Designs under three-way blocking with non-uniform clumps

This class of designs is developed for all natural numbers ≥ 5. First column is written in decreasing order mod v. Then develop each row cyclically mod v in increasing order. Now, the first clump is formed by taking the first (v-1) positions in the first row and 1 position in the second row. Again, second clump is formed by combining next (v-2) positions in the second row and 2 positions in third row.

Case 1: In case of even number of treatments, continue the above process till $\left(\frac{v}{2}-1\right)$ rows and take the remaining positions as the $\left(\frac{v}{2}\right)$ region so that the first $\left(\frac{v}{2}\right)$ rows are completely exhausted. The remaining rows are formed in a similar manner starting from the last row in reverse direction.

Case 2: For odd number of treatments, continue the above process till $\frac{(v-1)}{2}$ rows and take the remaining positions as the $[(v-1)/2]^{th}$ clump so that the first [(v-1)/2] rows are completely exhausted. Now, the v^{th} clump is formed by taking the (v-1) positions in the last row and 1 position in the second last row in a reverse direction. Again, $(v-1)^{th}$ clump is formed by combining next (v-2) positions in the $(v-1)^{th}$ row and 2 positions in $(v-2)^{th}$ row. Continue this process to obtain [(v+1)/2] clumps belonging to the bottom [(v+1)/2] rows.

In totality, v clumps are formed and shape of the clumps is different giving rise to irregular designs. For this class of designs, the parameters are v = r = p = q = m. The information matrix for estimating the elementary contrast pertaining to treatment effects is obtained for this class of designs using the Model 1 as $C_\tau = v\,(I_v - \frac{J_v}{v})$

These designs are variance balanced. The canonical efficiency factor of these designs in comparison to an orthogonal design having the same number of treatments and assuming the same variance has been computed and is unity for all the designs in this class. Parameters of the designs along with estimated variance are listed in Table 2 for v < 25.

Example 2: Let v = 6. The first column is written in decreasing order. Develop each row in increasing order cyclically mod 6. Regions are formed by the above said method. Regions are complete but shape of the regions is irregular.

6	1	2	3	4	5
5	6	1	2	3	4
4	5	6	1	2	3
3	4	5	6	1	2
2	3	4	5	6	1
1	2	3	4	5	6

For this design the parameters are v = r = p = q = m = 6. The information matrix for estimating the elementary contrast pertaining to treatment effects is of the $C_\tau = 6(I_6 - \frac{J_6}{6})$ form and the estimated variance of elementary contrasts pertaining to treatment effects is 0.3333.

Example 3: Let v = 7. The first column is written in decreasing order. Develop each row in increasing order cyclically mod 6. Clumps are formed by the above said method for odd numbers.

7	1	2	3	4	5	6
6	7	1	2	3	4	5
5	6	7	1	2	3	4
4	5	6	7	1	2	3
3	4	5	6	7	1	2
2	3	4	5	6	7	1
1	2	3	4	5	6	7

For this design the parameters are v = r = p = q = m = 7. The information matrix for estimating the elementary contrast pertaining to treatment effects is of the form $C_\tau = 7(I_7 - \frac{J_7}{7})$ and the estimated variance of elementary contrasts pertaining to treatment effects is 0.2857.

Table 2. List of designs

S. No.	v	r	p	q	m	V_{ij}
1	5	5	5	5	5	0.4000
2	6	6	6	6	6	0.3333
3	7	7	7	7	7	0.2857
4	8	8	8	8	8	0.2500
5	9	9	9	9	9	0.2220
6	10	10	10	10	10	0.2000
7	11	11	11	11	11	0.1818
8	12	12	12	12	12	0.1667
9	13	13	13	13	13	0.1538
10	14	14	14	14	14	0.1429
11	15	15	15	15	15	0.1333
12	16	16	16	16	16	0.1250
13	17	17	17	17	17	0.1176
14	18	18	18	18	18	0.1111
15	19	19	19	19	19	0.1053
16	20	20	20	20	20	0.1000

Designs under three-way blocking structure with non-uniform clumps having an empty node in each clump

First obtain a structurally complete design for v treatments under three-way blocking using above method in v rows, v columns and v clumps. In this arrangement, all of the main diagonal elements receive the same treatment v. Leaving the main diagonal empty, a structurally incomplete design for (v-1) treatments in v rows, v columns, v clumps where each treatment is repeated v times is obtained for irregular clumps. This method constructs designs for all natural numbers ≥ 5. The general form of information matrix for treatment effects is obtained as given below:

$$C_\tau = v(I_{v-1} - \frac{J_{v-1}}{v-1})$$

Designs were variance balanced. The canonical efficiency factor of these designs in comparison to an orthogonal design having the same number of treatments and assuming the same variance is computed and is unity for all the designs in this class. Parameters of the designs along with estimated variance are listed in Table 3 for $v < 25$.

Example 4: Let $v = 6$. A structurally complete design is obtained by using above method. Total number of clumps formed equal to 6. Leaving the main diagonal positions empty, a design under three-way blocking involving irregular clumps each clump having an empty node is obtained for 5 treatments:

*	1	2	3	4	5
5	*	1	2	3	4
4	5	*	1	2	3
3	4	5	*	1	2
2	3	4	5	*	1
1	2	3	4	5	*

For this design, the other parameters are $p = q = r = m = 6$. This design is variance balanced with efficiency factor as unity. The information matrix for estimating the elementary contrasts pertaining to treatment effects is of the form $C_\tau = 6(I_5 - \frac{J_5}{5})$ and the estimated variance of elementary contrasts pertaining to treatment effects is 0.4000.

Designs under three-way blocking structure having non-uniform clumps with an empty clump

Construct a Latin square for v (odd number) treatments such that positions falling on back-diagonal should receive distinct treatments. For obtaining such a Latin square, follow the steps given here:

First fill the entries of the first column allotting all odd treatments in ascending order and then the remaining positions are filled in by the even numbers in ascending order. Develop the other columns cyclically by adding one to the entry in the previous column and taking mod v, wherever required. In the resultant array, all the entries in the back diagonal are distinct.

Table 3. List of designs

S. No.	v	r	p	q	m	V_{ij}
1	5	4	5	5	5	0.4000
2	6	5	6	6	6	0.3333
3	7	6	7	7	7	0.2857
4	8	7	8	8	8	0.2500
5	9	8	9	9	9	0.2220
6	10	9	10	10	10	0.2000
7	11	10	11	11	11	0.1818
8	12	11	12	12	12	0.1667
9	13	12	13	13	13	0.1538
10	14	13	14	14	14	0.1429
11	15	14	15	15	15	0.1333
12	16	15	16	16	16	0.1250
13	17	16	17	17	17	0.1176
14	18	17	18	18	18	0.1111
15	19	18	19	19	19	0.1053
16	20	19	20	20	20	0.1000

Diagonal positions parallel to the back diagonal also receive distinct treatments. Leave this back diagonal empty. Clumps are formed by considering positions parallel to back diagonal together. There are v parallel diagonals in the backward direction including the one which is left blank. Thus, the final design has v rows, v columns, v clumps in which one complete clump is empty and each treatment is replicated (v-1) times. The designs are variance balanced.

The general form of information matrix for treatment effects is obtained utilizing the Model 1, which is of the form

$$C_\tau = \frac{v(v-3)}{(v-2)}[I_v - \frac{J_v}{v}]$$

The estimated variances of elementary contrasts pertaining to treatment effects have been computed using the SAS code developed for the purpose. The canonical efficiency factor of these designs in comparison to an orthogonal design having the same number of treatments and assuming the same variance is also computed. Parameters of the designs along with the estimated variances and efficiency factor are listed in Table 4 for $v < 25$.

Example 5: For $v = 7$, obtain a Latin square with distinct back diagonal entries as described above. Back diagonal is kept empty and hence, each treatment is replicated 6 times in the design. The design for $v = p = q = m = 7$ and $r = 6$:

1	2	3	4	5	6	*
3	4	5	6	7	*	2
5	6	7	1	*	3	4
7	1	2	*	4	5	6
2	3	*	5	6	7	1
4	*	6	7	1	2	3
*	7	1	2	3	4	5

The information matrix for this design is $\mathbf{C}_\tau = 5.6\mathbf{I}_7 - 0.8\mathbf{J}_7$ which shows that the design is variance balanced. The efficiency factor is computed as 0.9333 for the above design and this efficiency increases as v increases.

Illustration

An experiment was conducted to evaluate the effect of 6 feeds on the body weight of a group of animals of different ages and body weight put under six different housing conditions. The hypothetical data given below represents the body weight of 36 animals obtained after one month. Here, rows represent animals of different age groups (A_1, A_2, A_3, A_4, A_5 and A_6), columns represent animals having different housing conditions (H_1, H_2, H_3, H_4, H_5 and H_6) and clumps represent animals of different initial body

Table 4. List of designs

S. No.	v	r	p	q	m	V_{ij}	E_{can}
1	5	4	5	5	5	0.6000	0.8333
2	7	6	7	7	7	0.3571	0.9333
3	9	8	9	9	9	0.2592	0.9643
4	11	10	11	11	11	0.2045	0.9778
5	13	12	13	13	13	0.1692	0.9848
6	15	14	15	15	15	0.1444	0.9890
7	17	16	17	17	17	0.1260	0.9917
8	19	18	19	19	19	0.1118	0.9935
9	21	20	21	21	21	0.1005	0.9947
10	23	22	23	23	23	0.0913	0.9957

		Housing conditions					
		H_1	H_2	H_3	H_4	H_5	H_6
Age groups	A_1	31.50 [C_1] (6)	52.50 [C_1] (1)	58.09 [C_1] (2)	44.69 [C_1] (3)	63.75 [C_1] (4)	35.00 [C_3] (5)
	A_2	31.50 [C_1] (5)	49.36 [C_2] (6)	50.27 [C_2] (1)	59.41 [C_2] (2)	72.32 [C_2] (3)	47.25 [C_3] (4)
	A_3	52.50 [C_2] (4)	48.75 [C_2] (5)	46.41 [C_3] (6)	50.78 [C_3] (1)	94.96 [C_3] (2)	48.13 [C_3] (3)
	A_4	52.94 [C_4] (3)	80.44 [C_4] (4)	45.38 [C_4] (5)	50.27 [C_4] (6)	80.35 [C_5] (1)	62.56 [C_5] (2)
	A_5	68.25 [C_4] (2)	80.44 [C_5] (3)	74.25 [C_5] (4)	53.63 [C_5] (5)	78.89 [C_5] (6)	52.50 [C_6] (1)
	A_6	48.75 [C_4] (1)	95.06 [C_6] (2)	73.73 [C_6] (3)	87.14 [C_6] (4)	55.25 [C_6] (5)	51.19 [C_6] (6)

weight (C_1, C_2, C_3, C_4, C_5 and C_6). The numbers given in parenthesis are the six different feeds given to the animals.

Sas code for analyzing this data as per model (1) is given below:

SAS code

```
Data three_way_blocking;
Input age_gp house clump feed body_wt;
Cards;
1 1 1 6 31.50
1 2 1 1 31.50
.
.
.
6 5 6 5 52.50
6 6 6 6 51.19
;
PROC glm;
Class age_gp house clump feed;
Model body_wt = age_gp house clump feed /ss3;
Run;
quit;
```

Analysis of Variance (ANOVA) table revealed that all the effects are significant. It indicates the importance of accounting for the variability due to three sources viz., age_gp, house and clump for precise comparison among feeds.

RESULTS AND DISCUSSION

Four series of structurally complete/ incomplete designs under three-way blocking structure for irregular clumps have been developed for different parametric combinations. These designs are suitable when third source of variation does not permit blocking to be done in a regular pattern. First two series give structurally complete clumps while the other two give structurally incomplete designs under irregular clumps. First method yields designs having clumps with regular size while all other methods yield designs having non-uniform clumps. The first method is restricted to prime number of treatments, second and third methods are for all natural numbers ≥ 5 and the last method is for odd number of treatments. All the designs obtained are variance balanced and the elementary contrasts of treatment effects are estimated with constant variance. List of designs with different parametric combinations is given after each method of construction. Canonical efficiency factor for each design was worked out by writing SAS code in PROC IML and all the designs ere found to be highly efficient for sufficiently larger values of number of treatments.

ACKNOWLEDGMENT

The authors are grateful to the editor and the referee for their constructive comments which helped in improving the quality of the manuscript.

REFERENCES

Bailey R A, Kunert J and Martin R J. 1991. Some comments on Gerechte designs II. Randomization, analysis, and other methods that allow for inter-plot dependence. *Journal of Agronomy and Crop Science* **166**: 101–11.

Bailey R A, Cameron P J and Connelly R. 2008. Sudoku, Gerechte designs, resolutions, affine space, spreads, reguli, and hamming codes. *American Mathematical Monthly* **115**: 383–404.

Hui-Dong M O and Ru-Gen X U. 2008. Sudoku square - a new design in field experiment. *Acta Agronomica Sinica* **34**(9): 1489–93.

Saba F A and Sinha B K. 2014. SuDoKu as an experimental design - beyond the traditional Latin square design. *Statistics and Applications* **12** (1 and 2): 15–20.

Sarkar J and Sinha B K. 2014. Variants of SuDoKu as useful experimental designs. *Statistics and Applications* **12** (1 and 2): 35–60.

Subramani J and Ponnuswamy K N. 2009. Construction and analysis of Sudoku designs. *Model Assisted Statistics and Applications* **4**: 287–301.