ROW-COLUMN DESIGNS

Seema Jaggi
I.A.S.R.I., Library Avenue, New Delhi-110 012
seema@iasri.res.in

1. Introduction
Block designs are used when the heterogeneity present in the experimental material is in one-direction. However, when the heterogeneity present in the experimental material is in two directions i.e. rows and columns, then double grouping is done. The purpose of double grouping is to eliminate from the errors all differences among rows and equally all differences among columns. Under these situations, the experimental material should be arranged and the experiment conducted so that the differences among rows and columns represent major sources of variation that are known or suspected.

Example 1.1: In an animal experiment with the object of comparing the effects of four feeds, let young calves be the experimental units with their growth rate during a certain period as the variate under study. Let it be intended to eliminate the variation due to breeds and ages of the calves. So breed and age are the two factors that correspond to two sources of variation. The calves are, therefore, to come from, say, four breeds and four age groups. The 16 calves required for the experiment should each belong to a different breed-age combination. There should be 4 calves belonging to each breed and each of these 4 calves should come from a different age group. Both the factors of variation should be related to the variate under study so that their variability may influence the variability of the variate under study i.e. these are actually the controlled factors.

Example 1.2: Youden gave another application in greenhouse experiments on tobacco mosaic virus. The experimental unit was a single leaf, and the data consisted of the number of lesions produced per leaf by rubbing the leaf with a solution, which contained the virus. The numbers of lesions had been found to depend much more on inherent qualities of the plant than on the position of the plant on the greenhouse bench. Consequently, each row of the design was a single plant, so that the large differences in responsiveness, which existed among plants, did not contribute to the experimental errors. The columns were the positions, from top to bottom, of the five leaves, which were used on each plant. Since there was a fairly consistent gradient in responsiveness down each plant, this control also proved effective.

Designs used for the above situations are termed as Row-Column designs or designs for two-way elimination of heterogeneity. There can be incomplete rows and columns, complete rows and incomplete columns, incomplete rows and complete columns, complete rows and columns.

In a block design setting, the only source of variation, besides treatment, is the blocks. However, in the row-column setting, the sources of variation, besides the treatments, are the rows and column effects. If we ignore the rows and take columns as blocks, we get a block design with columns as blocks. Similarly ignoring columns and considering rows as blocks.
also results in a block design. One additional source of variation helps in reducing the experimental error and improving the precision of estimation.

The simplest and most commonly used row-column design with complete rows and columns is a **Latin Square design**. It is a design in which both rows and columns form a Randomized block design (RBD) i.e. treating rows as blocks forms a RBD and treating columns as blocks forms a RBD.

In example 1.2 above, if the number of leaves is sufficient so that every treatment can be applied to one leaf of a plant, then a Latin-square design could be used. But if the number of treatments is more than the number of leaves position available, then we must have incomplete columns. A row-column experimental plan with 7 treatments arranged in 4 rows and 7 columns, each treatment replicated 4 times is shown below.

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

- The row labels in the design are randomly assigned to the levels of the first blocking factor.
- The column labels in the design are randomly assigned to the levels of the second blocking factor.
- The treatment labels in the design are randomly assigned to the levels of the treatment factor.

Since there is only one experimental unit in each cell, there is no need for random assignment of experimental units to treatment labels within a cell.

### 2. General Row-Column Setting

In a general row-column setting, let the design have \( v \) treatments arranged in a \( p \times q \) array consisting of \( p \) rows and \( q \) columns. For this setting, three types of classification are possible.

1. Treatment Vs. row classification
2. Treatment Vs. column classification
3. Row Vs. column classification

It is taken that all the three classifications are possibly non-orthogonal. Let \( y_{ij(m)} \) be the response of the experimental unit occurring in the \( i^{th} \) row and \( j^{th} \) column to which the \( m^{th} \) treatment is applied. The model of response is

\[
y_{ij(m)} = \mu + \rho_i + c_j + \tau_m + e_{ij(m)} \quad (i)
\]

where \( \mu \) is the general mean, \( \rho_i \) is the \( i^{th} \) row effect, \( c_j \) is the \( j^{th} \) column effect, \( \tau_m \) is the \( m^{th} \) treatment effect and \( e_{ij(m)} \) are random errors assumed to be independently normally distributed with mean zero and variance \( \sigma^2 \). Here since there are three classificatory variables, therefore the design can be visualized to three block designs as given in (1) to (3) above. Therefore the information matrix \((C)\) for a row-column design is a function of the information matrix of the corresponding three block designs.
**Definition:** A treatment connected row-column design is called **Variance Balanced (VB) design** if and only if all the normalized treatment contrasts are estimated through the design with same variance. In other words, all the diagonal elements of the $C$ matrix are equal and all its off-diagonal elements are equal. A latin square design is a VB design.

**Youden Square design**
A Youden square design (YSD) is a design with incomplete columns by means of which two sources of variation can be eliminated. The rows of a YSD form a RBD and the columns form a Balanced Incomplete Block (BIB) design. These are basically symmetrical BIB designs by which the block to block variation can be eliminated. The $k$ units in each block can be thought of occupying $k$ different positions. With the help of YSD, the effects of such positions can also be eliminated.

**Definition:** A YSD is an arrangement of $v$ treatments in a $k \times v$ rectangular array such that every symbol occurs exactly once in each row and the columns form a symmetrical BIB design with parameters $v = b$, $r = k$, $\lambda$.

The $k$ units in each block of a symmetrical BIB design can always be so arranged that each treatment occupies each position once in some block or the other. When a symmetrical BIB design is obtained by developing an initial block such an arrangement is evident.

The position effects are evidently orthogonal to the treatments as each treatment occurs once in each position. The position effects are similarly orthogonal to the blocks as well. Hence, the analysis of YSD is the same as that of BIB designs excepting that a component of the sum of squares obtained from the position totals of the observations is subtracted from the error sum of squares of the BIB design.

Youden (1940) used these designs for the first time for green house studies. Following is a YSD with 7 treatments arranged in 3 rows and 7 columns.

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

Smith and Hartley (1948) proved the possibility of obtaining YSD from symmetrical BIB designs. The Youden squares can also be looked upon as an incomplete $v \times v$ latin square with $k$ suitably chosen columns. If a column or a row is omitted from a latin square, the resultant design is always a YSD.

The model of response for a YSD is the same as in (a) with $i=1,2,\ldots,k$: $j=1,2,\ldots,v$ and $m=1,2,\ldots,v$. Let

- $R_1$, $R_2,\ldots, R_k$ be the $k$ row totals;
- $C_1,C_2,\ldots, C_v$ be the $v$ column totals,
- $T_1$, $T_2,\ldots, T_v$ be the $v$ treatment totals;
- $G$ be the grand total.

Further let $Q_m$ be the adjusted $m^{th}$ treatment total obtained by subtracting the sum of column means in which $m^{th}$ treatment occurs from $T_m$ ($m = 1,2,\ldots,v$). Further let $\mathbf{Q} = (Q_1, Q_2,\ldots, Q_m)'$.

The information matrix pertaining to the estimation of treatment effects for a YSD is
\[ C = \frac{\lambda v}{k} \left[ I - 11'/v \right] = \frac{(k-1)v}{v-1} \left[ I - 11'/v \right]. \]

Therefore a YSD is a VB design. Further

\[ \hat{\tau}_m = \frac{k}{\lambda v} Q_m, m = 1, 2, \ldots, v. \]

and the Analysis of variance table is as given below.

<table>
<thead>
<tr>
<th>Sources of variation</th>
<th>Degrees of Freedom (D.F.)</th>
<th>Sum of Squares (S.S.)</th>
<th>Mean sum of squares (M.S.)</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rows</td>
<td>k-1</td>
<td>( R'R \left( \frac{G^2}{v} \right) )</td>
<td>MS_r</td>
<td>MS_r/MS_e</td>
</tr>
<tr>
<td>Columns</td>
<td>v-1</td>
<td>( C'C \left( \frac{G^2}{k} \right) )</td>
<td>MS_c</td>
<td>MS_c/MS_e</td>
</tr>
<tr>
<td>Treatments (eliminating rows and columns)</td>
<td>v-1</td>
<td>( Q'C - Q = \frac{k}{\lambda v} \left( Q'Q \right) )</td>
<td>MS_t</td>
<td>MS_t/MS_e</td>
</tr>
<tr>
<td>Error</td>
<td>( vk-2v-k+2 )</td>
<td>By subtraction*</td>
<td>MS_e</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>( vk-1 )</td>
<td>( y'y - \frac{G^2}{vk} )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Error S.S. = Total S.S. - Rows S.S. - Columns S.S. - Treatments (eliminating rows and columns) S.S.

The significance of the treatment effects can be tested by comparing the F value of table with \( F_{v-1, vk-2v-k+2, \alpha} \). The calculation of treatment S.S. is identical to the one for a BIB design and the other calculations like Rows S.S, Columns S.S are the routine ones. Analogous to the results on BIB design, the C.D. for comparing any two treatment effects is

\[ (t_{vk - 2v-k+2, \alpha/2}) \sqrt{2k MS_e / \lambda v} \]

**Illustration:** In one of the experiments, the experimenter is interested in making comparisons among 7 treatments and there are 28 experimental units available. These 28 experimental units are arranged in a Youden Square design with 4 rows and 7 columns with one observation per cell. The parameters of the design are \( v \) (number of treatments) = 7, \( p \) (number of rows) = 4, \( q \) (number of columns) =7, \( r \) (replication of treatments) = 4. The layout of the design along with the observations is given below.

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2(4.00)</td>
<td>3(5.30)</td>
<td>4(1.10)</td>
<td>5(16.90)</td>
<td>6(16.90)</td>
<td>7(10.30)</td>
<td>1(294.00)</td>
<td></td>
</tr>
<tr>
<td>7(17.50)</td>
<td>1(220.00)</td>
<td>2(12.20)</td>
<td>3(15.50)</td>
<td>4(11.00)</td>
<td>5(26.50)</td>
<td>6(27.20)</td>
<td></td>
</tr>
<tr>
<td>6(37.00)</td>
<td>7(26.00)</td>
<td>1(310.00)</td>
<td>2(22.70)</td>
<td>3(24.20)</td>
<td>4(21.40)</td>
<td>5(31.30)</td>
<td></td>
</tr>
<tr>
<td>5(46.80)</td>
<td>6(44.20)</td>
<td>7(34.30)</td>
<td>1(282.00)</td>
<td>2(33.70)</td>
<td>3(33.70)</td>
<td>4(30.50)</td>
<td></td>
</tr>
</tbody>
</table>

The analysis of variance for the above design for testing the equality of treatment effects is obtained as follows:
From the above ANOVA table for the row-column setting, it may be seen that the probability of getting an F value greater than 143.56 is as small as 0.0001. The treatments effect, therefore, differ significantly at 0.01 percent level of significance.

In the above Youden square design if the rows are ignored and the columns are treated as blocks, then the following analysis of variance is obtained.

\[ R^2 = 0.9736, \quad CV = 33.9323, \quad Root \ Mean \ Square \ Error = 20.0601 \]

A comparison of MSE in the two ANOVA tables indicates that there has been considerable reduction in the MSE by forming rows and columns.

For testing the rows and columns effect, we perform the following analysis of variance to get adjusted sum of squares.

\[ R^2 = 0.9736, \ CV = 33.9323, \ Root \ Mean \ Square \ Error = 20.0601 \]

The row effects are highly significant while the column effects are not significantly different. In this example, therefore, the formation of rows and columns has been very effective and has resulted into a considerable reduction in the error mean square.

**Cyclic Row-Column Designs:** A cyclic row-column design is a row-column design in which the columns form a cyclic block design and the rows are complete blocks. The class of cyclic row-column designs is very large and a design with \( q = v \) columns can always be found when a Youden design does not exist. For example, consider an experiment for comparing \( v=8 \) treatments when there are two blocking factors having \( p=3 \), \( v=q=8 \) because there does
not exist a BIB design for 8 treatments in 8 blocks of size 3. A cyclic design that may be suitable for the experiment is

\[
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
2 & 3 & 4 & 5 & 6 & 7 & 8 & 1 \\
4 & 5 & 6 & 7 & 8 & 1 & 2 & 3 \\
\end{array}
\]

Here treatment pairs (1,5), (2,6), (3,7) and (4,8) never occur together in a column, but all other pairs of treatments occur in exactly one column.

Pearce (1952) developed a number of designs of the latin square type. He considered the analysis and layout for latin squares with an additional column, with an additional row and an additional column, and with a column added and a row omitted. Shrikhande (1951) considers some general classes of designs for the two-way elimination of heterogeneity.

Kiefer (1975) introduced **Generalized Youden Designs (GYD)** as a generalization of usual latin square and Youden square designs. A \( p \times q \) row-column design in \( v \) treatments is called a GYD if it is a balanced block design in both directions. Clearly, for a GYD, \( r_1 = r_2 = \ldots = r_v = \frac{pq}{v} \). The \( C \)-matrix of a GYD has all the diagonal elements equal and all the off-diagonal elements equal. In other words, a GYD is a variance-balanced design. When \( p = q = v \), a GYD reduces to a latin square design, and for \( p < v \), \( q = v(mv) \), a GYD is simply a Youden square design.

When number of rows is equal to the number of columns in a row-column design, Cheng (1981) introduced another important class of variance balanced designs called **Pseudo-Youden designs (PYD)**. A PYD is a row-column design with number of rows equal to the number of columns and the rows and the columns taken together form a balanced block design. Clearly, for a PYD, \( r_1 = r_2 = \ldots = r_v = \frac{p^2}{v} \). A square GYD is a PYD, but the converse is not always true.

A \( p \times q \) array containing entries from a finite set of \( v \) treatment symbols will be called a Youden Type (YT) design if the \( i^{th} \) symbol occurs in each row \( m_i \) times, \( i = 1, 2, \ldots, v \) where \( m_i = \frac{r_i}{p} \) and \( r_i \) is the replication of the \( i^{th} \) treatment in the array. YSD’s are covered in this class of designs.

**References**


