UNIFORMITY TRIAL - SIZE AND SHAPE OF PLOT

Krishan Lal
I.A.S.R.I., Library Avenue, New Delhi - 11 0012

In agricultural experimentation every research worker is interested in ascertaining the relative worth of a set of treatments with reasonable confidence. The simple procedure of trying these treatments each in a different field or plot does not seem to adequate to ascertain their relative worth with reasonable confidence. For even after discovering from such a trail that some treatments have given a better performance than others the experimenter is left wondering whether the differences observed are due to treatments, to inherent fertility differences in the soil or some other accidental factors. Ideally the research worker would like to try the treatments under identical conditions but even with the most uniform land that he can select, he finds that the inherent variation in the soil is quite considerable and the simple procedure of trying out different treatments on single plots side by side in the same field does not suffice for assessing the intrinsic worth of the treatments. A good idea of the nature and extent of fertility variation in land can be obtained from the results of what are known as *uniformity trials*. In other words uniformity trial is planned to determine suitable size and shape of the plot and the number of plots in a block.

Uniformity trial involves growing a particular crop on a field or piece of land with uniform conditions. All sources of variation except that due to a native soil differences, are kept constant. At the time of harvest the entire field is divided into smaller units of same size and shape and the produce from each such units is recorded separately. The smallest the basic units, the more detailed is the measurement of soil heterogeneity. In the past, large number of research workers has attempted to study the soil fertility variation through various methods. Some of the methods for soil fertility variation/plot size are given below:

1. Fertility Contour Map

An approach to describe the heterogeneity of land is to construct the fertility contour map. This is constructed by taking the moving averages of yields of unit plots and demarcating the regions of same fertility by considering those areas, which have yield of same magnitude. This approach of describing the variation in fertility has been adopted by large number of workers in India and abroad.

2. Maximum Curvature Method

In this method basic units of uniformity trials are combined to form new units. The new units are formed by combining columns, rows or both. Combination of columns and rows be done in such a way that no columns or rows is left out. For each set of units, the coefficient of variation (CV) is computed. A curve is plotted by taking the plot size (in terms of basic units) on X-axis and the CV values on the Y-axis of graph sheet. The point at which the curve takes a turn, *i.e.*, the point of maximum curvature is located by inspection. The value corresponding to the point of maximum curvature will be optimum plot size.

Harris (1915, 1920) has shown that adjacent areas are correlated, as such the hypothesis of no correlation is not tenable. He utilizes these criteria for subdividing the field into uniform areas. He suggested use of intra-class correlation as a measure of heterogeneity. If this correlation coefficient is in the neighbourhood of zero then field could be considered as homogenous field and whatever plot size is adopted, it will not lead to a large experimental error. These correlation coefficients do not give any idea of plot size.

3. Fairfield Smith's Variance Law

Keeping in view of drawbacks of various methods given above, Smith (1938) gave a empirical relations between variance and plot size. He developed an empirical model representing the relationship between plot size and variance of mean per plot. This model is given by the equation

$$V_x = \frac{V_1}{x^b} \text{ or } \log V_x = \log V_1 - b \log x.$$

where x is number of basic units in a plot, V_x is the variance of mean per plot of x units, V_1 is the variance of mean per plot of one unit, and b is the characteristics of soil and measure of correlation among contiguous units if $b=1$, $V_x = \frac{V_1}{x}$ and the units making up the plots of x unit are not correlated at all. On the other hand, if $b=0$, the x units are perfectly correlated and $V_x = V_1$ so there is no gain due to the larger size of plot. In general, b will be between 0 and 1 so that the larger plot gives more information with the same number of plots. In that case, larger area for the purpose of experiment will be used. The values of V_1 and b are determined by the principle of least squares.

Example:

Table-1 Grain yield (g/m^2) of Rice Variety IR8 from Uniformity Test covering an area 18x 36 m.

	Column							
Row	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	842	844	808	822	979	954	965	906
2	803	841	870	970	943	914	916	836
3	773	782	860	822	932	971	765	875
4	912	887	815	937	844	661	841	844
5	874	792	803	793	818	799	767	855
6	908	875	899	788	867	790	831	757
7	875	907	921	963	875	880	898	802
8	891	928	871	875	865	777	738	796
9	823	784	754	873	764	775	752	753
10	785	794	764	822	714	748	724	717
11	785	808	823	826	801	712	826	665
12	829	895	774	891	841	815	834	778
13	861	883	739	762	725	717	746	766
14	906	885	790	655	690	769	765	719
15	819	911	788	654	742	786	791	779
16	893	862	769	727	725	721	739	736
17	813	750	742	872	746	812	705	724
18	816	758	811	702	728	741	757	732
19	676	783	734	626	782	704	782	707
20	813	809	695	707	753	680	720	683

	1	2	3	4	5	6	7	8
21	801	764	701	716	753	680	706	665
22	718	784	730	750	733	705	728	667
23	756	725	821	685	681	738	630	599
24	789	681	732	669	681	698	689	622
25	652	622	695	677	698	666	691	688
26	729	650	700	764	680	681	645	622
27	698	713	714	734	651	649	675	614
28	745	677	685	711	688	614	585	534
29	964	727	648	664	623	629	616	594
30	671	729	690	687	705	622	523	526
31	717	694	727	719	669	630	701	645
32	652	713	656	584	517	572	574	539
33	605	708	684	715	659	629	632	596
34	559	722	726	705	571	637	637	577
35	589	681	690	570	619	624	580	570
36	614	633	619	658	678	673	652	602

9	10	11	12	13	14	15	16	17	18
898	856	808	920	808	889	943	894	968	917
858	926	922	910	872	805	775	846	947	965
853	936	927	779	865	720	566	893	914	861
809	778	945	876	901	802	836	778	923	949
792	858	912	839	813	740	730	632	813	914
751	774	863	902	771	747	819	699	670	934
874	928	872	834	892	760	753	720	751	894
855	901	792	752	722	781	739	733	783	786
820	798	847	858	811	875	659	661	759	767
736	724	838	769	819	823	724	750	764	764
759	738	867	725	794	755	730	638	724	734
760	822	803	754	703	743	728	692	748	671
662	634	743	719	710	682	694	675	709	720
743	770	728	740	691	767	648	715	655	665
645	810	816	746	729	814	718	721	708	722
672	814	756	748	714	718	694	704	915	705
640	757	708	750	767	638	754	767	763	685
623	786	805	786	739	727	767	738	659	695
672	703	698	758	762	625	623	699	662	613
757	782	789	811	789	769	751	648	680	696
680	650	690	699	768	751	701	665	603	680
703	684	777	747	713	696	717	732	712	679
629	703	780	720	709	697	731	661	627	644
672	704	705	625	677	704	648	605	585	651
682	713	670	708	707	695	681	716	626	637
661	728	715	775	690	726	669	766	709	645
634	635	639	690	694	637	590	640	658	609
533	671	600	647	592	595	563	634	666	644
619	631	628	591	675	654	640	718	667	649
661	683	619	709	620	651	676	728	547	682
638	714	633	670	649	665	557	734	674	727
545	629	636	580	607	654	585	674	608	612
627	644	661	682	690	636	665	731	753	640
561	590	646	639	672	636	651	684	584	622
568	589	550	622	623	706	725	738	669	636
590	605	538	682	651	653	680	696	633	660

Smith's Index of Soil Heterogeneity

Step-1 Combine the *r x c* basic units to simulate plots of different sizes and shapes. Use only combinations that fit exactly into the whole area, i.e. the product of simulated plots and the number of basic units per plot must equal to the total number of basic units.

Step-2 For each of the simulated plots constructed in Step-1, compute the yield total *T* as the sum of basic units to construct that plot and compute $V_{(x)}$, $V_{x.}$

$$V_{(x)} = \frac{\sum_{i=1}^{w} T_i^2}{x} - \frac{G^2}{rc}, \text{ where } w = rc/x \text{ is the total number of plots simulated.}$$

The between plot variance for plot of size 2 x 1 m is

$$V_{(2)} = \frac{\sum_{i=1}^{324} T_i^2}{2} - \frac{G^2}{36x18} = 31{,}370 \text{ and}$$

$$V_x = \frac{V_{(x)}}{x^2}, \qquad V_2 = \frac{V_{(2)}}{2^2} = 7{,}842.$$

Step-3 For each plot size having more than one shape, test the homogeneity of between-plot variance $V_{(x)}$, to determine the significance of plot-orientation (plot shape) effect, by using F test or the Chi-square test. For each plot size whose plot shape effect is non-significant, compute the average of V_x values over all plot shapes. For others, use the lowest value.

For plot of size 2 m^2 there are two shapes only. So, F = 31,370/31,309 = 1.00 (NS).

Step-4 Using the values of the variance per unit are V_x computed in Step 3, estimate the regression coefficient between V_x and plot size x. We fit the equation

$$V_x = \frac{V_1}{x^b} \text{ or } \log V_x = \log V_1 - b \log x \text{ or } Y = cx$$

where $Y = \log V_x - \log V_1$, $c = -b$ and $X = \log x$

$$c = \frac{\sum_{i=1}^{m} w_i x_i Y_i}{\sum_{i=1}^{m} w_i x_i^2}, \quad \sum_{i=1}^{12} w_i x_i Y_i = -3.8503 \text{ and } \sum_{i=1}^{12} w_i x_i^2 = 27.9743$$

c = -0.1376 or b=0.1376.

Here w_I is the number of plot shapes used in computing the average variance per unit area of the i^{th} plot and *m* is the total number of plots of different sizes.

Thus the Fairfield Smith's equation is $\hat{V}_x = \frac{9{,}041}{x^{0.1376}}$.

Table-2 Between-plot variance [$V_{(x)}$], variance per unit area (V_x) and coefficient of variability (CV) of plots of various sizes and shapes, calculated from Rice Uniformity Data in Table-1

Plot Size and Shape						
Size m^2	Width m	Length m	Plot numbers	$V_{(x)}$	V_x	CV %
1	1	1	648	9,041	9,041	13.0
2	2	1	324	31,370	7,842	12.1
3	3	1	216	66,396	7,377	11.7
6	6	1	108	235,112	6,531	11.0
9	9	1	72	494,497	6,105	10.7
2	1	2	324	31,301	7,827	12.1
4	2	2	162	114,515	7,157	11.5
6	3	2	108	247,140	6,865	11.3
12	6	2	54	908,174	6,307	10.8
18	9	2	36	1,928,177	5,951	10.5
3	1	3	216	66,330	7,370	11.7
6	2	3	108	247,657	6,879	11.3
9	3	3	72	537,201	6,632	11.1
18	6	3	36	1,981,408	6,115	10.7
27	9	3	24	4,231,622	5,805	10.4
4	1	4	162	113,272	7,080	11.5
8	2	4	81	427,709	6,683	11.1
12	3	4	54	943,047	6,549	11.0
24	6	4	27	3,526,179	6,121	10.7
36	9	4	18	7,586,647	5,854	10.4
6	1	6	108	238,384	6,622	11.1
12	2	6	54	913,966	6,347	10.9
18	3	6	36	2,021,308	6,239	10.8
36	6	6	18	7,757,823	5,986	10.5
9	1	9	72	514,710	6,354	10.9
18	2	9	36	2,017,537	6,227	10.8
27	3	9	24	4,513,900	6,192	10.7

This law can further be used for arriving at an optimum plot size. He has recommended the cost function $C = C_1+C_2 X$ where C_1 = overhead cost which is independent of plot size and C_2 is the consideration of cost by a unit increase in the plot size. Optimum value of plot size is one which minimises the cost per unit of information viz. $(C_1+C_2 X)$

Assuming that variance is given by

$$X_{opt} = \frac{b}{1-b}\frac{C_1}{C_2}$$

The optimum plot size had been worked out for different cost ratio and for values of b.

Optimum Plot Size

The optimum or recommended size of an experimental unit (plot) cannot be given without first considering several factors:

i) Practical Consideration: Certain practical aspects may dictate the size of experimental unit. In animal experiments the pen's or cages may be already constructed and not easily changed. The pasture or paddocks size may be

determined by those already available and fenced. If grain combined and other power equipment are used, a fairly large plot may be essential. In green house studies, the experimental unit may have to be small. Experimental resources available would also determine the plot size.

ii) Nature of experimental material : The plot size is different for oat and for corn, pen or cages size for chicken and cattle are also different.

iii) Number of treatment per block or per incomplete block: For large number treatments to be tested, an incomplete blocks design may be used.

iv) Variability among individual or units within the experimental unit (V_s) relative to variability among experimental unit (V_p) treated alike. The variance of treatment mean is proportional to $[V_p + V_s/k]$. The relative size of the V_p and V_s has considerable effects on optimum size.

v) Cost: Let C_s be cost of an individual item within experimental unit which is independent of cost of experimental units. C_p be the cost of experimental unit, independent of individuals in the unit. Then the cost per treatment with a single replication is $kC_s + C_p = C_t$ and the total cost of experiment is a random variable C_t (v treatment, r replications) the optimum size thus depends on the ratio of C_s and C_p.

Shape of Plots

Cochran (1940) has also considered the problem of shape of plots for various types of fields. His results can very well understand by the following example. Suppose we have v=9 treatment to compare and we wish to select the plot shape with the smallest average experimental error variance when the direction of fertility gradient is unknown. The two extreme shapes selected are rectangular [plan (a)] and Square [plan(b)] as shown in Figure A. Let us consider the Case for which a linear fertility, gradient exists. Suppose it is parallel to AB (Fig. A) so that plots lie perpendicular to the gradient. The sum of squares among the 9 plots would be $8\sigma_i^2 + 60g^2$. If fertility to the gradient was parallel to AC, the plots would be parallel to the gradient and all the plots would be equally affected by the gradient. In this case the sum of squares of the 9 plots would be $8\sigma_I^2$

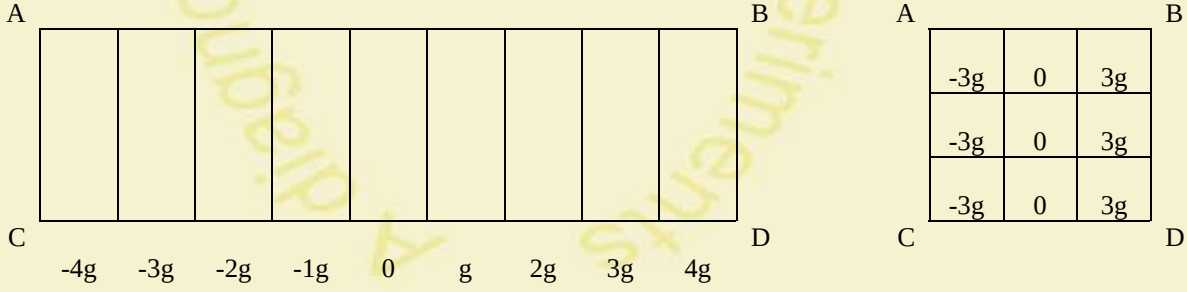


Fig. A

If σ_I^2 is the random variance within blocks, independent of shape of plot, the mean squares as affected by the two plot shapes, rectangular and square as shown in Fig. A are given below:

Fertility Gradient	d f	Average value of Mean Square	
		Plan (a)	Plan (b)
Parallel to AC	8	σ_I^2	$\sigma_I^2 + (54/8)g^2$
Parallel to AB	8	$\sigma_I^2 +(60/8)g^2$	$\sigma_I^2 +(54/8)g^2$
Average	8	$\sigma_I^2 +(30/8)g^2$	$\sigma_I^2 +(54/8)g^2$

If the plots were square as in plan (b) the sum of squares among plots would be the same for both cases. The average mean square for long narrow plots is smaller than that of the square plots.

As in the experimental design, the plots are generally arranged within blocks. Therefore, for the efficient planning the information on the efficiency of different block size is also of great importance. For working out the relative efficiency block sizes the ratio of error variance of a particular block arrangement to that without block arrangement can be worked out. This ratio is expressed as percentage and was taken as efficiency for that block arrangement.

In many a uniformity trials reported, it is observed that while the Fairfield Smith law explains the relationship between the plot size and average variation (or coefficient of variation) very well, the variation for different shapes of the same plot are not of the same order. In such situation, while the law can be used for arriving optimum plot size, the shape of the plot need to be arrived after examining the variation associated with the shapes. Studies have revealed that the relationship between the block size (for a fixed plot size and shape) and variation also follow a similar law viz. $Y=ax^b$ where Y is the variation (coefficient of variation) and x is the block size.

The repeated analysis of uniformity trial data by super imposing different sizes and shape of plots and blocks and studying the variances or coefficient of variation can be work out and studied with the help of the relationship to arrive the optimum plot/block size/shape. This data can also be used to study the relative efficiency of various experimental design like Completely Randomised Design, Randomised Block Design, Incomplete Block Design, Confounded Factorials, Latin Square.

References

Federer, W. T. (1967) *Experimental design: Theory and Application.* Oxford & IBH Publishing Co., New Delhi.

Gomez, K. A. and Gomez, A. A. (1984) *A statistical procedure for agricultural research.* John Wiley & Sons, New York.

Smith, M.F.C. 1938. An empirical law describing heterogeneity in the yields of agricultural crops. *J. Agril. Sci. Comb.*, **28**, 1-23.

Chapter 4

Experimental Designs and Their Analysis

Design of experiment means how to design an experiment in the sense that how the observations or measurements should be obtained to answer a query in a valid, efficient and economical way. The designing of the experiment and the analysis of obtained data are inseparable. If the experiment is designed properly keeping in mind the question, then the data generated is valid and proper analysis of data provides the valid statistical inferences. If the experiment is not well designed, the validity of the statistical inferences is questionable and may be invalid.

It is important to understand first the basic terminologies used in the experimental design.

Experimental unit:

For conducting an experiment, the experimental material is divided into smaller parts and each part is referred to as an experimental unit. The experimental unit is randomly assigned to treatment is the experimental unit. The phrase "randomly assigned" is very important in this definition.

Experiment:

A way of getting an answer to a question which the experimenter wants to know.

Treatment

Different objects or procedures which are to be compared in an experiment are called treatments.

Sampling unit:

The object that is measured in an experiment is called the sampling unit. This may be different from the experimental unit.

Factor:

A factor is a variable defining a categorization. A factor can be fixed or random in nature. A factor is termed as a fixed factor if all the levels of interest are included in the experiment.

A factor is termed as a random factor if all the levels of interest are not included in the experiment and those that are can be considered to be randomly chosen from all the levels of interest.

Replication:

It is the repetition of the experimental situation by replicating the experimental unit.

Experimental error:

The unexplained random part of the variation in any experiment is termed as experimental error. An estimate of experimental error can be obtained by replication.

Treatment design:

A treatment design is the manner in which the levels of treatments are arranged in an experiment.

Example: (Ref.: Statistical Design, G. Casella, Chapman and Hall, 2008)

Suppose some varieties of fish food is to be investigated on some species of fishes. The food is placed in the water tanks containing the fishes. The response is the increase in the weight of fish. The experimental unit is the tank, as the treatment is applied to the tank, not to the fish. Note that if the experimenter had taken the fish in hand and placed the food in the mouth of fish, then the fish would have been the experimental unit as long as each of the fish got an independent scoop of food.

Design of experiment:

One of the main objectives of designing an experiment is how to verify the hypothesis in an efficient and economical way. In the contest of the null hypothesis of equality of several means of normal populations having the same variances, the analysis of variance technique can be used. Note that such techniques are based on certain statistical assumptions. If these assumptions are violated, the outcome of the test of a hypothesis then may also be faulty and the analysis of data may be meaningless. So the main question is how to obtain the data such that the assumptions are met and the data is readily available for the application of tools like analysis of variance. The designing of such a mechanism to obtain such data is achieved by the design of the experiment. After obtaining the sufficient experimental unit, the treatments are allocated to the experimental units in a random fashion. Design of experiment provides a method by which the treatments are placed at random on the experimental units in such a way that the responses are estimated with the utmost precision possible.

Principles of experimental design:

There are three basic principles of design which were developed by Sir Ronald A. Fisher.

(i) Randomization

(ii) Replication

(iii) Local control

(i) Randomization

The principle of randomization involves the allocation of treatment to experimental units at random to avoid any bias in the experiment resulting from the influence of some extraneous unknown factor that may affect the experiment. In the development of analysis of variance, we assume that the errors are random and independent. In turn, the observations also become random. The principle of randomization ensures this.

The random assignment of experimental units to treatments results in the following outcomes.

a) It eliminates systematic bias.
b) It is needed to obtain a representative sample from the population.
c) It helps in distributing the unknown variation due to confounded variables throughout the experiment and breaks the confounding influence.

Randomization forms a basis of a valid experiment but replication is also needed for the validity of the experiment.

If the randomization process is such that every experimental unit has an equal chance of receiving each treatment, it is called **complete randomization**.

(ii) Replication:

In the replication principle, any treatment is repeated a number of times to obtain a valid and more reliable estimate than which is possible with one observation only. Replication provides an efficient way of increasing the precision of an experiment. The precision increases with the increase in the number of observations. Replication provides more observations when the same treatment is used, so it increases precision. For example, if the variance of x is σ^2 than variance of the sample mean $\overline{x}$ based on n observation is $\frac{\sigma^2}{n}$. So as n increases, $Var(\overline{x})$ decreases.

(ii) Local control (error control)

The replication is used with local control to reduce the experimental error. For example, if the experimental units are divided into different groups such that they are homogeneous within the blocks, then the variation among the blocks is eliminated and ideally, the error component will contain the variation due to the treatments only. This will, in turn, increase the efficiency.

Complete and incomplete block designs:

In most of the experiments, the available experimental units are grouped into blocks having more or less identical characteristics to remove the blocking effect from the experimental error. Such design is termed as **block designs**.

The number of experimental units in a block is called the **block size**.
If
size of block = number of treatments
and
each treatment in each block is randomly allocated,
then it is a **full replication** and the design is called a **complete block design**.

In case, the number of treatments is so large that a full replication in each block makes it too heterogeneous with respect to the characteristic under study, then smaller but homogeneous blocks can be used. In such a case, the blocks do not contain a full replicate of the treatments. Experimental designs with blocks containing an incomplete replication of the treatments are called **incomplete block designs**.

Completely randomized design (CRD)

The CRD is the simplest design. Suppose there are v treatments to be compared.

- All experimental units are considered the same and no division or grouping among them exist.
- In CRD, the v treatments are allocated randomly to the whole set of experimental units, without making any effort to group the experimental units in any way for more homogeneity.
- Design is entirely flexible in the sense that any number of treatments or replications may be used.
- The number of replications for different treatments need not be equal and may vary from treatment to treatment depending on the knowledge (if any) on the variability of the observations on individual treatments as well as on the accuracy required for the estimate of individual treatment effect.

Example: Suppose there are 4 treatments and 20 experimental units, then

- the treatment 1 is replicated, say 3 times and is given to 3 experimental units,
- the treatment 2 is replicated, say 5 times and is given to 5 experimental units,
- the treatment 3 is replicated, say 6 times and is given to 6 experimental units

and

- finally, the treatment 4 is replicated [20-(6+5+3)=]6 times and is given to the remaining 6 experimental units.

- All the variability among the experimental units goes into experimented error.
- CRD is used when the experimental material is homogeneous.
- CRD is often inefficient.
- CRD is more useful when the experiments are conducted inside the lab.
- CRD is well suited for the small number of treatments and for the homogeneous experimental material.

Layout of CRD

Following steps are needed to design a CRD:

- Divide the entire experimental material or area into a number of experimental units, say n.
- Fix the number of replications for different treatments in advance (for given total number of available experimental units).
- No local control measure is provided as such except that the error variance can be reduced by choosing a homogeneous set of experimental units.

Procedure

Let the v treatments are numbered from 1,2,...,v and n_i be the number of replications required for i^{th} treatment such that $\sum_{i=1}^{v} n_i = n.$

- Select n_1 units out of n units randomly and apply treatment 1 to these n_1 units.

  (**Note**: This is how the randomization principle is utilized is CRD.)
- Select n_2 units out of $(n-n_1)$ units randomly and apply treatment 2 to these n_2 units.
- Continue with this procedure until all the treatments have been utilized.
- Generally, the equal number of treatments are allocated to all the experimental units unless no practical limitation dictates or some treatments are more variable or/and of more interest.

Analysis

There is only one factor which is affecting the outcome – treatment effect. So the set-up of one-way analysis of variance is to be used.

y_{ij} : Individual measurement of j^{th} experimental units for i^{th} treatment $i = 1,2,...,v$, $j = 1,2,...,n_i$.

y_{ij} : Independently distributed following $N(\mu+\alpha_i,\sigma^2)$ with $\sum_{i=1}^{v} n_i\alpha_i = 0$.

μ : overall mean

α_i : i^{th} treatment effect

$H_0 : \alpha_1 = \alpha_2 = ... = \alpha_v = 0$

$H_1 :$ All $\alpha_i's$ are not equal.

The data set is arranged as follows:

Treatments			
1	2	...	v
y_{11}	y_{21}	...	y_{v1}
y_{12}	y_{22}	...	y_{v2}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
y_{1n_1}	y_{2n_2}	...	y_{vn_v}
T_1	T_2	...	T_v

where $T_i = \sum_{j=1}^{n_i} y_{ij}$ is the treatment total due to i^{th} effect,

$G = \sum_{i=1}^{v} T_i = \sum_{i=1}^{v} \sum_{j=1}^{n_i} y_{ij}$ is the grand total of all the observations.

In order to derive the test for H_0, we can use either the likelihood ratio test or the principle of least squares. Since the likelihood ratio test has already been derived earlier, so we choose to demonstrate the use of the least-squares principle.

The linear model under consideration is

$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad i = 1,2,...,v, \; j = 1,2,...,n_i$

where $\varepsilon_{ij}'s$ are identically and independently distributed random errors with mean 0 and variance σ^2. The normality assumption of $\varepsilon's$ is not needed for the estimation of parameters but will be needed for deriving the distribution of various involved statistics and in deriving the test statistics.

Let $S = \sum_{i=1}^{v} \sum_{j=1}^{n_i} \varepsilon_{ij}^2 = \sum_{i=1}^{v} \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i)^2.$

Minimizing S with respect to μ and α_i, the normal equations are obtained as

$$\frac{\partial S}{\partial \mu} = 0 \Rightarrow n\mu + \sum_{i=1}^{v} n_i \alpha_i = 0$$

$$\frac{\partial S}{\partial \alpha_i} = 0 \Rightarrow n_i \mu + n_i \alpha_i = \sum_{j=1}^{n_i} y_{ij} \;\; i = 1,2,..,v.$$

Solving them using $\sum_{i=1}^{v} n_i \alpha_i = 0$, we get

$\hat{\mu} = \bar{y}_{oo}$

$\hat{\alpha}_i = \bar{y}_{io} - \bar{y}_{oo}$

where $\bar{y}_{io} = \frac{1}{n_i}\sum_{j=1}^{n_i} \bar{y}_{ij}$ is the mean of observation receiving the i^{th} treatment and $\bar{y}_{oo} = \frac{1}{n}\sum_{i=1}^{v}\sum_{j=1}^{n_i} \bar{y}_{ij}$ is the mean of all the observations.

The fitted model is obtained after substituting the estimate $\hat{\mu}$ and $\hat{\alpha}_i$ in the linear model. Using the fitted model, we can write

$$y_{ij} = \bar{y}_{oo} + (\bar{y}_{io} - \bar{y}_{oo}) + (y_{ij} - \bar{y}_{io})$$

or $(y_{ij} - \bar{y}_{oo}) = (\bar{y}_{io} - \bar{y}_{oo}) + (y_{ij} - \bar{y}).$

Squaring both sides and summing over all the observation, we have

$$\sum_{i=1}^{v}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{oo})^2 = \sum_{i=1}^{v} n_i(\bar{y}_{io} - \bar{y}_{oo})^2 + \sum_{i=1}^{v}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{io})^2$$

$$\downarrow \qquad\qquad \downarrow \qquad\qquad \downarrow$$

or $$\begin{pmatrix}\text{Total sum}\\ \text{of squares}\end{pmatrix} = \begin{pmatrix}\text{Sum of squares}\\ \text{due to treatment}\\ \text{effects}\end{pmatrix} + \begin{pmatrix}\text{Sum of squares}\\ \text{due to error}\end{pmatrix}$$

or $$TSS = SSTr + SSE$$

- Since $\sum_{i=1}^{v}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{oo}) = 0,$ so TSS is based on the sum of $(n-1)$ squared quantities. The TSS carries only $(n-1)$ degrees of freedom.

- Since $\sum_{i=1}^{v} n_i(\bar{y}_{io} - \bar{y}_{oo}) = 0,$ so $SSTr$ is based only on the sum of $(v\text{ -}1)$ squared quantities. The $SSTr$ carries only $(v\text{ -}1)$ degrees of freedom.

- Since $\sum_{i=1}^{n_i} n_i(\bar{y}_{ij} - \bar{y}_{io}) = 0$ for all $i = 1,2,...,v,$ so SSE is based on the sum of squaring n quantities like $(y_{ij} - \bar{y}_{io})$ with v constraints $\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{io}) = 0,$ So SSE carries $(n - v)$ degrees of freedom.

- Using the Fisher-Cochran theorem,

  $TSS = SSTr + SSE$

  with degrees of freedom partitioned as

  $(n - 1) = (v\text{ - }1) + (n - v).$

Moreover, equality in $TSS = SSTr + SSE$ has to hold exactly. To ensure that the equality holds exactly, we find one of the sums of squares through subtraction. Generally, it is recommended to find SSE by subtraction as

$$SSE = TSS - SSTr$$

$$TSS = \sum_{i=1}^{v}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{io})^2$$

$$= \sum_{i=1}^{v}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{G^2}{n}$$

where

$$G = \sum_{i=1}^{v}\sum_{j=1}^{n_i} y_{ij}.$$

$$SSTr = \sum_{j=1}^{n_i} n_i(\bar{y}_{io} - \bar{y}_{oo})^2$$

$$= \sum_{i=1}^{v}\left(\frac{T_i^2}{n_i}\right) - \frac{G^2}{n}$$

where $T_i = \sum_{j=1}^{n_i} y_{ij}$

$\frac{G^2}{n}$: correction factor .

Now under $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_v = 0$, the model become

$$Y_{ij} = \mu + \varepsilon_{ij},$$

and minimizing $S = \sum_{i=1}^{v}\sum_{j=1}^{n_i}\varepsilon_{ij}^2$

with respect to μ gives

$$\frac{\partial S}{\partial \mu} = 0 \Rightarrow \hat{\mu} = \frac{G}{n} = \bar{y}_{oo}.$$

The SSE under H_0 becomes

$$SSE = \sum_{i=1}^{v}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{oo})^2$$

and thus $TSS = SSE$. This TSS under H_0 contains the variation only due to the random error whereas the earlier $TSS = SSTr + SSE$ contains the variation due to treatments and errors both. The difference between the two will provides the effect of treatments in terms of the sum of squares as

$$SSTr = \sum_{i=1}^{v} n_i(\bar{y}_i - \bar{y}_{oo})^2$$

Distributions and decision rules:

Using the normal distribution property of ε_{ij}'s, we find that y_{ij}'s are also normal as they are the linear combination of ε_{ij}'s.

- $\frac{SSTr}{\sigma^2} \sim \chi^2(v-1)$ under H_0
- $\frac{SSE}{\sigma^2} \sim \chi^2(n-v)$ under H_0
- $SSTr$ and SSE are independently distributed
- $\frac{MStr}{MSE} \sim F(v-1, n-v)$ under H_0.
- Reject H_0 at α^* level of significance if $F > F_{\alpha^*;v-1,n-v.}$

[Note: We denote the level of significance here by α^* because α has been used for denoting the factor]

The analysis of variance table is as follows

Source of variation	Degrees of freedom	Sum of squares	Mean sum of squares	F
Between treatments	v - 1	$SSTr$	$MSTr$	$\frac{MSTr}{MSE}$
Errors	n - v	SSE	MSE	
Total	n - 1	TSS		

Expectations

$$
\begin{aligned}
E(SSE) &= \sum_{i=1}^{v}\sum_{j=1}^{n_i} E(y_{ij} - y_{io})^2 \\
&= \sum_{i=1}^{v}\sum_{j=1}^{n_i} (\varepsilon_{ij} - \bar{\varepsilon}_{io})^2 \\
&= \sum_{i=1}^{v}\sum_{j=1}^{n_i} E(\varepsilon_{ij}^2) - \sum_{i=1}^{v} n_i E(\bar{\varepsilon}_{io.}^2) \\
&= n\sigma^2 - \sum_{i=1}^{v} n_i \frac{\sigma^2}{n_i} \\
&= (n-v)\sigma^2
\end{aligned}
$$

$$
E(MSE) = E\left(\frac{SSE}{n-v}\right) = \sigma^2
$$

$$
\begin{aligned}
E(SSTr) &= \sum_{i=1}^{v} n_i E(\bar{y}_{io} - \bar{y}_{oo})^2 \\
&= \sum_{i=1}^{v} n_i E(\alpha_i + \bar{\varepsilon}_{io} - \bar{\varepsilon}_{oo})^2 \\
&= \sum_{i=1}^{v} n_i \alpha_i^2 + \left[\sum_{i=1}^{v} n_i \bar{\varepsilon}_{io}^2 - n\bar{\varepsilon}_{oo}^2 \right] \\
&= \sum_{i=1}^{v} n_i \alpha_i^2 + \left[\sum_{i=1}^{v} n_i \frac{\sigma^2}{n_i} - n\frac{\sigma^2}{n} \right] \\
&= \sum_{i=1}^{v} n_i \alpha_i^2 + (v-1)\sigma^2
\end{aligned}
$$

$$
E(MSTr) = E\left(\frac{SStr}{v-1}\right) = \frac{1}{v-1}\sum_{i=1}^{v} n_i \alpha_i^2 + \sigma^2.
$$

In general $E\left(MSTr\right) \neq \sigma^2$ but under H_0, all $\alpha_i = 0$ and so $E(MSTr) = \sigma^2$.

Randomized Block Design

If a large number of treatments are to be compared, then a large number of experimental units are required. This will increase the variation among the responses and CRD may not be appropriate to use. In such a case when the experimental material is not homogeneous and there are *v* treatments to be compared, then it may be possible to

- group the experimental material into blocks of sizes *v* units.
- Blocks are constructed such that the experimental units within a block are relatively homogeneous and resemble to each other more closely than the units in the different blocks.
- If there are *b* such blocks, we say that the blocks are at *b* levels. Similarly, if there are *v* treatments, we say that the treatments are at *v* levels. The responses from the *b* levels of blocks and *v* levels of treatments can be arranged in a two-way layout. The observed data set is arranged as follows:

		Treatments (Factor B)						Block totals
		1	2		j		v	
Blocks (Factor A)	1	y_{11}	y_{12}	…	y_{1j}	…	y_{1v}	B_1
	2	y_{21}	y_{22}	…	y_{2j}	…	y_{2v}	B_2
	. . .	. . .	. . .		. . .		. . .	. . .
	i	y_{i1}	y_{i2}	…	y_{ij}	…	y_{iv}	B_i
	. . .		. . .		. . .		. . .	. . .
	b	y_{b1}	y_{b2}	…	y_{bj}	…	y_{bv}	B_b
Treatment totals		T_1	T_2	…	T_j	…	T_v	Grand total (G)

Layout:

A two-way layout is called a randomized block design (RBD) or a randomized complete block design (RCB) if, within each block, the v treatments are randomly assigned to v experimental units such that each of the $v!$ ways of assigning the treatments to the units has the same probability of being adopted in the experiment and the assignment in different blocks are statistically independent.

The RBD utilizes the principles of design - randomization, replication and local control - in the following way:

1. Randomization:

- Number the v treatments $1,2,\ldots,v$.
- Number the units in each block as $1, 2,\ldots,v$.
- Randomly allocate the v treatments to v experimental units in each block.

2. Replication

Since each treatment is appearing in each block, so every treatment will appear in all the blocks. So each treatment can be considered as if replicated the number of times as the number of blocks. Thus in RBD, the number of blocks and the number of replications are same.

3. Local control

Local control is adopted in RBD in the following way:

- First form the homogeneous blocks of the experimental units.
- Then allocate each treatment randomly in each block.

The error variance now will be smaller because of homogeneous blocks and some variance will be parted away from the error variance due to the difference among the blocks.

Example:

Suppose there are 7 treatments denoted as $T_1, T_2, .., T_7$ corresponding to 7 levels of a factor to be included in 4 blocks. So one possible layout of the assignment of 7 treatments to 4 different blocks in an RBD is as follows

Block 1	T_2	T_7	T_3	T_5	T_1	T_4	T_6
Block 2	T_1	T_6	T_7	T_4	T_5	T_3	T_2
Block 3	T_7	T_5	T_1	T_6	T_4	T_2	T_3
Block 4	T_4	T_1	T_5	T_6	T_2	T_7	T_3

Analysis

Let

y_{ij} : Individual measurements of j^{th} treatment in i^{th} block, $i = 1,2,...,b,\ j = 1,2,...,v$.

y_{ij}'s are independently distributed following $N(\mu + \beta_i + \tau_j, \sigma^2)$

where μ: overall mean effect

β_i: i^{th} block effect

τ_j: j^{th} treatment effect

such that $\sum_{i=1}^{b} \beta_i = 0,\ \sum_{j=1}^{v} \tau_j = 0$.

There are two null hypotheses to be tested.

- related to the block effects

$$H_{0B}: \beta_1 = \beta_2 = ... = \beta_b = 0.$$

- related to the treatment effects

$$H_{0T}: \tau_1 = \tau_2 = ... = \tau_v = 0.$$

The linear model, in this case, is a two-way model as

$$y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij},\ i = 1,2,..,b;\ j = 1,2,..,v$$

where ε_{ij} are identically and independently distributed random errors following a normal distribution with mean 0 and variance σ^2.

The tests of hypothesis can be derived using the likelihood ratio test or the principle of least squares. The use of likelihood ratio test has already been demonstrated earlier, so we now use the principle of least squares.

$$\text{Minimizing } S = \sum_{i=1}^{b}\sum_{j=1}^{v} \varepsilon_{ij}^2 = \sum_{i=1}^{b}\sum_{j=1}^{v} (y_{ij} - \mu - \beta_i - \tau_j)^2$$

and solving the normal equation

$$\frac{\partial S}{\partial \mu} = 0,\ \frac{\partial S}{\partial \beta_i} = 0,\ \frac{\partial S}{\partial \tau_j} = 0 \text{ for all } i = 1,2,..,b,\ j = 1,2,..,v.$$

the least squares estimators are obtained as

$$\hat{\mu} = \bar{y}_{oo},$$
$$\hat{\beta}_i = \bar{y}_{io} - \bar{y}_{oo},$$
$$\hat{\tau}_j = \bar{y}_{oj} - \bar{y}_{oo}.$$

Using the fitted model (obtained after substituting the estimated values of the parameters in the model), we can write

$$y_{ij} = \bar{y}_{oo} + (\bar{y}_{io} - \bar{y}_{oo}) + (\bar{y}_{oj} - \bar{y}_{oo}) + (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo}).$$

Squaring both sides and summing over *i* and *j* gives

$$\sum_{i=1}^{b}\sum_{j=1}^{v} (\bar{y}_{ij} - \bar{y}_{oo})^2 = v\sum_{i=1}^{b} (\bar{y}_{io} - \bar{y}_{oo})^2 + b\sum_{j=1}^{v} (\bar{y}_{oj} - \bar{y}_{oo})^2 + \sum_{i=1}^{b}\sum_{j=1}^{v} (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})^2$$

or $TSS = SSBl + SSTr + SSE$

with degrees of freedom partitioned as

$$bv - 1 = (b-1) + (v-1) + (b-1)(v-1).$$

The reason for the number of degrees of freedom for different sums of squares is the same as in the case of CRD.

Here $TSS = \sum_{i=1}^{b}\sum_{j=1}^{v}(y_{ij} - \bar{y}_{oo})^2$

$$= \sum_{i=1}^{b}\sum_{j=1}^{v} y_{ij}^2 - \frac{G^2}{bv}$$

$\frac{G^2}{bv}$: correction factor.

$G = \sum_{i=1}^{b}\sum_{j=1}^{v} y_{ij}$: Grand total of all the observation.

$$SSBl = v\sum_{i=1}^{b}(\bar{y}_{io} - \bar{y}_{oo})^2$$

$$= \sum_{i=1}^{b}\frac{B_i^2}{v} - \frac{G^2}{bv}$$

$B_i = \sum_{j=1}^{v} y_{ij} : i^{th}$ block total

$$SSTr = b\sum_{j=1}^{v}(\bar{y}_{oj} - \bar{y}_{oo})^2$$

$$= \sum_{j=1}^{v}\frac{T_j^2}{b} - \frac{G^2}{bv}$$

$T_j = \sum_{i=1}^{b} y_{ij} : j^{th}$ treatment total

$$SSE = \sum_{i=1}^{b}\sum_{j=1}^{v}(y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})^2.$$

The expectations of mean squares are

$$E(MSBl) = E\left(\frac{SSBl}{b-1}\right) = \sigma^2 + \frac{v}{b-1}\sum_{i=1}^{b}\beta_i^2$$

$$E(MSTr) = E\left(\frac{SSTr}{v-1}\right) = \sigma^2 + \frac{b}{v-1}\sum_{j=1}^{v}\tau_j^2$$

$$E(MSE) = E\left(\frac{SSE}{(b-1)(v-1)}\right) = \sigma^2.$$

Moreover,

$$(b-1)\frac{SSBl}{\sigma^2} \sim \chi^2(b-1)$$

$$(v-1)\frac{SSTr}{\sigma^2} \sim \chi^2(v-1)$$

$$(b-1)(v-1)\frac{SSE}{\sigma^2} \sim \chi^2(b-1)(v-1).$$

Under $H_{0B} : \beta_1 = \beta_2 = ... = \beta_b = 0,$

$$E(MSBl) = E(MSE)$$

and $SSBl$ and SSE are independent , so

$$F_{bl} = \frac{MSBl}{MSE} \sim F((b-1,(b-1)(v-1)).$$

Similarly, under $H_{0T} : \tau_1 = \tau_2 = ... = \tau_v = 0,$ so

$$E(MSTr) = E(MSE)$$

and $SSTr$ and SSE are independent, so

$$F_{Tr} = \frac{MSTr}{MSE} \sim F(v-1),(b-1)(v-1)).$$

Reject H_{0B} *if* $F_{be} > F_{\alpha}((b-1),(b-1)(v-1)$

Reject H_{0T} *if* $F_{Tr} > F_{\alpha}((v-1),(b-1)(v-1))$

If H_{0B} is accepted, then it indicates that the blocking is not necessary for future experimentation.

If H_{0T} is rejected then it indicates that the treatments are different. Then the multiple comparison tests are used to divide the entire set of treatments into different subgroup such that the treatments in the same subgroup have the same treatment effect and those in the different subgroups have different treatment effects.

The analysis of variance table is as follows

Source of variation	Degrees of freedom	Sum of squares	Mean squares	F
Blocks	b - 1	$SSBl$	$MSBl$	F_{Bl}
Treatments	v - 1	$SSTr$	$MSTr$	F_{Tr}
Errors	(b - 1)(v - 1)	SSE	MSE	
Total	bv - 1	TSS		

Latin Square Design

The treatments in the RBD are randomly assigned to b blocks such that each treatment must occur in each block rather than assigning them at random over the entire set of experimental units as in the CRD. There are only two factors – block and treatment effects – which are taken into account and the total number of experimental units needed for complete replication are bv where b and v are the numbers of blocks and treatments respectively.

If there are three factors and suppose there are b, v and k levels of each factor, then the total number of experimental units needed for a complete replication are bvk. This increases the cost of experimentation and the required number of experimental units over RBD.

In Latin square design (LSD), the experimental material is divided into rows and columns, each having the same number of experimental units which is equal to the number of treatments. The treatments are allocated to the rows and the columns such that each treatment occurs once and only once in each row and in each column.

In order to allocate the treatment to the experimental units in rows and columns, we take help from Latin squares.

Latin Square:

A Latin square of order p is an arrangement of p symbols in p^2 cells arranged in p rows and p columns such that each symbol occurs once and only once in each row and in each column. For example, to write a Latin square of order 4, choose four symbols – A, B, C and D. These letters are Latin letters which are used as symbols. Write them in a way such that each of the letters out of A, B, C and D occurs once and only once in each row and each column. For example, as

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

This is a Latin square.

We consider first the following example to illustrate how a Latin square is used to allocate the treatments and in getting the response.

Example:

Suppose different brands of petrol are to be compared with respect to the mileage per litre achieved in motor cars.

Important factors responsible for the variation in mileage are

- the difference between individual cars.
- the difference in the driving habits of drivers.

We have three factors – cars, drivers and petrol brands. Suppose we have

- 4 types of cars denoted as 1, 2, 3, 4.
- 4 drivers that are represented by a, b, c, d.
- 4 brands of petrol are indicated as A, B, C, D.

Now the complete replication will require $4 \times 4 \times 4 = 64$ the number of experiments. We choose only 16 experiments. To choose such 16 experiments, we take the help of the Latin square. Suppose we choose the following Latin square:

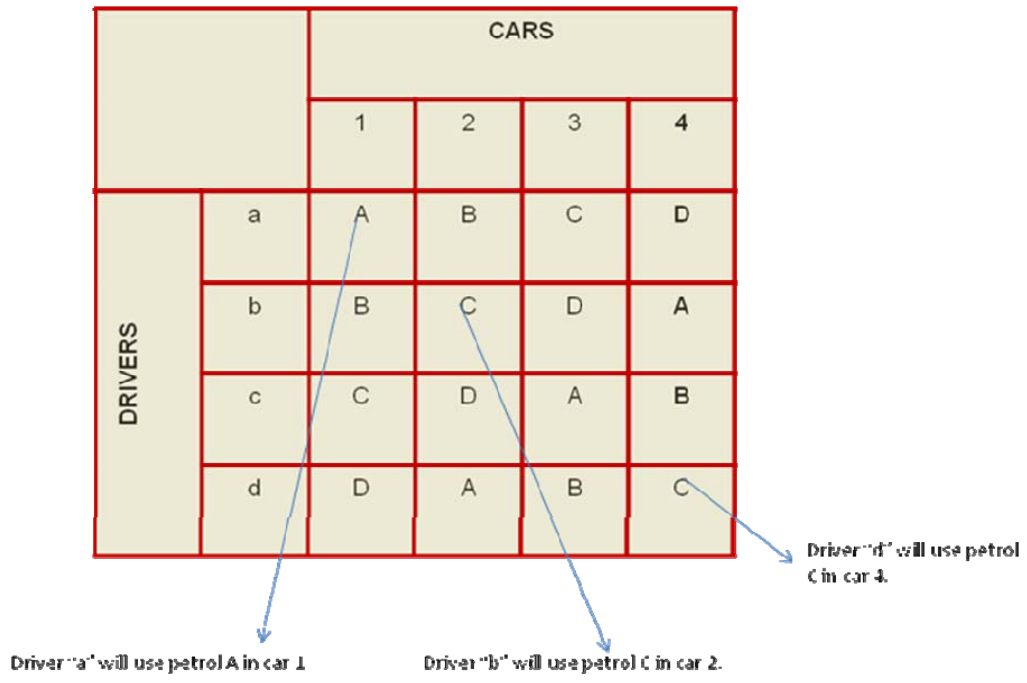
A B C D

B C D A

C D A B

D A B C

Write them in rows and columns and choose rows for drivers, columns for cars and letter for petrol brands. Thus 16 observations are recorded as per this plan of treatment combination (as shown in the next figure) and further analysis is carried out. Since such design is based on Latin square, so it is called as a Latin square design.



Another choice of a Latin square of order 4 is

C B A D
B C D A
A D C B
D A B C

This will again give a design different from the previous one. The 16 observations will be recorded again but based on different treatment combinations.

Since we use only 16 out of 64 possible observations, so it is an incomplete 3-way layout in which each of the 3 factors – cars, drivers and petrol brands are at 4 levels and the observations are recorded only on 16 of the 64 possible treatment combinations.

Thus in an LSD,

- the treatments are grouped into replication in two-ways
    - once in rows and
    - and in columns,
- rows and columns variations are eliminated from the within treatment variation.
    - In RBD, the experimental units are divided into homogeneous blocks according to the blocking factor. Hence it eliminates the difference among blocks from the experimental error.
    - In LSD, the experimental units are grouped according to two factors. Hence two effects (like as two block effects) are removed from the experimental error.
    - So the error variance can be considerably reduced in LSD.

The LSD is an incomplete three-way layout in which each of the three factors, viz, rows, columns and treatments, is at v levels each and observations only on v^2 of the v^3 possible treatment combinations are taken. Each treatment combination contains one level of each factor.

The analysis of data in an LSD is conditional in the sense it depends on which Latin square is used for allocating the treatments. If the Latin square changes, the conclusions may also change.

We note that Latin squares play an important role is an LSD, so first we study more about these Latin squares before describing the analysis of variance.

Standard form of Latin square

A Latin square is in the standard form if the symbols in the first row and first columns are in the **natural order** (Natural order means the order of alphabets like A, B, C, D,…).

Given a Latin square, it is possible to rearrange the columns so that the first row and first column remain in a natural order.

Example: Four standard forms of 4×4 Latin square are as follows.

A B C D	A B C D	A B C D	A B C D
B A D C	B C D A	B D A C	B A D C
C D B A	C D A B	C A D B	C D A B
D C A B	D A B C	D C B A	D C B A

For each standard Latin square of order p, the p rows can be permuted in $p!$ ways. Keeping a row fixed, vary and permute $(p - 1)$ columns in $(p - 1)!$ ways. So there are $p!(p - 1)!$ different Latin squares.

For illustration

Size of square	Number of Standard squares	Value of $p!(1 - p)!$	Total number of different squares
3 x 3	1	12	12
4 x 4	4	144	576
5 x 5	56	2880	161280
6 x 6	9408	86400	812851250

Conjugate:

Two standard Latin squares are called conjugate if the rows of one are the columns of other.

For example

A B C D		A B C D
B C D A	and	B C D A
C D A B		C D A B
D A B C		D A B C

are conjugate. In fact, they are **self conjugate**.

A Latin square is called **self conjugate** if its arrangement in rows and columns are the same.

Transformation set:

A set of all Latin squares obtained from a single Latin square by permuting its rows, columns and symbols is called a transformation set.

From a Latin square of order p, $p!(p-1)!$ different Latin squares can be obtained by making $p!$ permutations of columns and $(p-1)!$ permutations of rows which leaves the first row in place. Thus

Number of different Latin squares of order p in a transformation set = $p!(p-1)!$ X number of standard Latin squares in the set

Orthogonal Latin squares

If two Latin squares of the same order but with different symbols are such that when they are superimposed on each other, every ordered pair of symbols (different) occurs exactly once in the Latin square, then they are called orthogonal.

Graeco-Latin square:

A pair of orthogonal Latin squares, one with Latin symbols and the other with Greek symbols form a Graeco-Latin square.

For example

A B C D	$\alpha\ \beta\ \gamma\ \delta$
B A D C	$\delta\ \gamma\ \beta\ \alpha$
C D A B	$\beta\ \alpha\ \delta\ \gamma$
D C B A	$\gamma\ \delta\ \alpha\ \beta$

is a Graeco-Latin square of order 4.

Graeco Latin squares design enables to consider one more factor than the factors in Latin square design. For example, in the earlier example, if there are four drivers, four cars, four petrol and each petrol has four varieties, as α, β, γ and δ, then Graeco-Latin square helps in deciding the treatment combination as follows:

		Cars			
		1	2	3	4
Drivers	a	$A\alpha$	$B\beta$	$C\gamma$	$D\delta$
	b	$B\delta$	$A\gamma$	$D\beta$	$C\alpha$
	c	$C\beta$	$D\alpha$	$A\delta$	$B\gamma$
	d	$D\gamma$	$C\delta$	$B\alpha$	$A\beta$

Now

$A\alpha$ means: Driver 'a' will use the α variant of petrol A in Car 1.

$B\gamma$ means: Driver 'c' will use the γ variant of petrol B in Car 4

and so on.

Mutually orthogonal Latin square

A set of Latin squares of the same order is called a set of mutually orthogonal Latin square (or a hyper Graeco-Latin square) if every pair in the set is orthogonal. The total number of mutually orthogonal Latin squares of order p is at most (p - 1).

Analysis of LSD (one observation per cell)

In designing an LSD of order p,

- choose one Latin square at random from the set of all possible Latin squares of order p.
- Select a standard Latin square from the set of all standard Latin squares with equal probability.
- Randomize all the rows and columns as follows:
  - Choose a random number, less than p, say n_1 and then 2nd row is the $n_1{}^{th}$ row.
  - Choose another random number less than p, say n_2 and then 3rd row is the $n_2{}^{th}$ row and so on.
  - Then do the same for the column.
- For Latin squares of the order less than 5, fix the first row and then randomize rows and then randomize columns. In Latin squares of order 5 or more, need not to fix even the first row. Just randomize all rows and columns.

Example:

Suppose following Latin square is chosen

A B C D E

B C D E A

D E A B C

E A B C D

C D E A B

Now randomize rows, e.g., 3rd row becomes 5th row and 5th row becomes 3rd row . The Latin square becomes

A B C D E

B C D E A

C D E A B

E A B C D

D E A B C.

Now randomize columns, say 5th column becomes 1st column, 1st column becomes 4th column and 4th column becomes 5th column

E B C A D

A C D B E

D A B E C

C E A D B

B D E C A

Now use this Latin square for the assignment of treatments.

y_{ijk} : Observation on k^{th} treatment in i^{th} row and j^{th} block, $i = 1,2,...,v,\ j = 1,2,...,v,\ k = 1,2,...,v$.

Triplets (I, j, k) take on only the v^2 values indicated by the chosen particular Latin square selected for the experiment.

y_{ijk} 's are independently distributed as $N(\mu + \alpha_i + \beta_j + \tau_k, \sigma^2)$.

Linear model is

$$y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \varepsilon_{ijk}, i = 1,2,...,v; j = 1,2,...,v; k = 1,2,...,v$$

where ε_{ijk} are random errors which are identically and independently distributed following $N(0,\sigma^2)$.

with $\sum_{i=1}^{v} \alpha_i = 0,\ \sum_{j=1}^{v} \beta_i = 0,\ \sum_{k=1}^{v} \tau_k = 0,$

α_i : main effect of rows

β_j :main effect of columns

γ_k :main effect of treatments.

The null hypothesis under consideration are

$H_{0R} : \alpha_1 = \alpha_2 = ... = \alpha_v = 0$

$H_{0C} : \beta_1 = \beta_2 = ... = \beta_v = 0$

$H_{0T} : \tau_1 = \tau_2 = ... = \tau_v = 0$

The analysis of variance can be developed on the same lines as earlier.

Minimizing $S = \sum_{i=1}^{v}\sum_{j=1}^{v}\sum_{k=1}^{v} \varepsilon_{ijk}^2$ with respect to μ, α_i, β_j and τ_k given the least-squares estimate as

$\hat{\mu} = \bar{y}_{ooo}$

$\hat{\alpha}_i = \bar{y}_{ioo} - \bar{y}_{ooo} \qquad i = 1,2,...,v$

$\hat{\beta}_j = \bar{y}_{ojo} - \bar{y}_{ooo} \qquad j = 1,2,...,v$

$\hat{\tau}_k = \bar{y}_{ook} - \bar{y}_{ooo} \qquad k = 1,2,...,v.$

Using the fitted model based on these estimators, the total sum of squares can be partitioned into the mutually orthogonal sum of squares *SSR*, *SSC*, *SSTr* and *SSE* as

$TSS = SSR + SSC + SSTr + SSE$

where

TSS: Total sum of squares = $\sum_{i=1}^{v}\sum_{j=1}^{v}\sum_{k=1}^{v}(y_{ijk} - \bar{y}_{ooo})^2 = \sum_{i=1}^{v}\sum_{j=1}^{v}\sum_{k=1}^{v} y_{ijk}^2 - \frac{G^2}{v^2}$

SSR: Sum of squares due to rows = $v\sum_{i=1}^{v}(\bar{y}_{ioo} - \bar{y}_{ooo})^2 = \frac{\sum_{i=1}^{v} R_i^2}{v} - \frac{G^2}{v^2}$; where $R_i = \sum_{j=1}^{v}\sum_{k=1}^{v} y_{ijk}$

SSC: Sum of squares due to column = $v\sum_{j=1}^{v}(\bar{y}_{ojo} - \bar{y}_{ooo})^2 = \frac{\sum_{i=1}^{v} C_j^2}{v} - \frac{G^2}{v^2}$; where $C_j = \sum_{i=1}^{v}\sum_{k=1}^{v} y_{ijk}$

SSTr : Sum of squares due to treatment = $v\sum_{k=1}^{v}(\bar{y}_{ook} - \bar{y}_{ooo})^2 = \frac{\sum_{i=1}^{v} T_k^2}{v} - \frac{G^2}{v^2}$; where $T_k = \sum_{i=1}^{v}\sum_{j=1}^{v} y_{ijk}$

Degrees of freedom carried by *SSR, SSC* and *SSTr* are $(v - 1)$ each.

Degrees of freedom carried by *TSS* is $v^2 - 1$.

Degree of freedom carried by *SSE* is $(v - 1)(v - 2)$.

The expectations of mean squares are obtained as

$$E(MSR) = E\left(\frac{SSR}{v-1}\right) = \sigma + \frac{v}{v-1}\sum_{i=1}^{v}\alpha_i^2$$

$$E(MSC) = E\left(\frac{SSC}{v-1}\right) = \sigma^2 + \frac{v}{v-1}\sum_{j=1}^{v}\beta_j^2$$

$$E(MSTr) = E\left(\frac{SSTr}{v-1}\right) = \sigma^2 + \frac{v}{v-1}\sum_{k=1}^{v}\tau_k^2$$

$$E(MSE) = E\left(\frac{SSE}{(v-1)(v-2)}\right) = \sigma^2.$$

Thus

$$\text{- under } H_{0R},\ F_R = \frac{MSR}{MSE} \sim F((v-1),(v-1)(v-2))$$

$$\text{- under } H_{0C},\ F_C = \frac{MSC}{MSE} \sim F((v-1),(v-1)(v-2))$$

$$\text{- under } H_{0T},\ F_T = \frac{MSTr}{MSE} \sim F((v-1),(v-1)(v-2)).$$

Decision rules:

Reject H_{0R} at level α if $F_R > F_{1-\alpha;v(-1),(v-1)(v-2)}$

Reject H_{0C} at level α if $F_C > F_{1-\alpha;(v-1),(v-1)(v-2)}$

Reject H_{0T} at level α if $F_T > F_{1-\alpha;(v-1),(v-1)(v-2)}$.

If any null hypothesis is rejected, then use multiple comparison test.

The analysis of variance table is as follows

Source of variation	Degrees of freedom	Sum of squares	Mean sum of squares	F
Rows	v - 1	*SSR*	*MSR*	F_R
Columns	v - 1	*SSC*	*MSC*	F_C
Treatments	v - 1	*SSTr*	*MSTr*	F_T
Error	$(v - 1)(v - 2)$	*SSE*	*MSE*	
Total	$v^2 - 1$	*TSS*		

Missing plot techniques:

It happens many time in conducting the experiments that some observation are missed. This may happen due to several reasons. For example, in a clinical trial, suppose the readings of blood pressure are to be recorded after three days of giving the medicine to the patients. Suppose the medicine is given to 20 patients and one of the patients doesn't turn up for providing the blood pressure reading. Similarly, in an agricultural experiment, the seeds are sown and yields are to be recorded after few months. Suppose some cattle destroy the crop of any plot or the crop of any plot is destroyed due to storm, insects etc.

In such cases, one option is to

- somehow estimate the missing value on the basis of available data,
- replace it back in the data and make the data set complete.

Now conduct the statistical analysis on the basis of completed data set as if no value was missing by making necessary adjustments in the statistical tools to be applied. Such an area comes under the purview of "missing data models" and a lot of development has taken place. Several books on this issue have appeared, e.g.,

- Little, R.J.A. and Rubin, D.B. (2002). Statistical Analysis with Missing Data, 2nd edition, New York: John Wiley.
- Schafer, J.L. (1997). Analysis of Incomplete Multivariate Data. Chapman & Hall, London etc.

We discuss here the classical missing plot technique proposed by Yates which involve the following steps:

- Estimate the missing observations by the values which makes the error sum of squares to be minimum.
- Substitute the unknown values by the missing observations.
- Express the error sum of squares as a function of these unknown values.
- Minimize the error sum of squares using the principle of maxima/minima, i.e., differentiating it with respect to the missing value and put it to zero and form a linear equation.
- Form as much linear equation as the number of unknown values (i.e., differentiate the error sum of squares with respect to each unknown value).
- Solve all the linear equations simultaneously and solutions will provide the missing values.
- Impute the missing values with the estimated values and complete the data.
- Apply analysis of variance tools.
- The error sum of squares thus obtained is corrected but the treatment sum of squares is not corrected.
- The number of degrees of freedom associated with the total sum of squares is subtracted by the number of missing values and adjusted in the error sum of squares. No change in the degrees of freedom of sum of squares due to treatment is needed.

Missing observations in RBD

One missing observation:

Suppose one observation in $(i, j)^{th}$ cell is missing and let this be x. The arrangement of observations in RBD then will look like as follows:

		Treatments (Factor B)						Block totals
		1	**2**		j		v	
Blocks (Factor A)	**1**	y_{11}	y_{12}	…	y_{1j}	…	y_{1v}	$B_1 = y_{1o}$
	2	y_{21}	y_{22}	…	y_{2j}	…	y_{2v}	$B_2 = y_{2o}$
	. . .	. . .	. . .		. . .		. . .	. . .
	i	y_{i1}	y_{i2}	…	$y_{ij}=x$	…	y_{iv}	$B_i = y'_{io} + x$
	. . .		. . .		. . .		. . .	. . .
	b	y_{b1}	y_{b2}	…	y_{bj}	…	y_{bv}	B_b
Treatment totals		$T_1 = y_{o1}$	$T_2 = y_{o2}$	…	$T_j = y'_{oj} + x$	…	$T_v = y_{ob}$	**Grand total** $(G) = y'_{oo} + x$

where y'_{oo} : total of known observations

y'_{io} : total of known observations in i^{th} block

y'_{oj} : total of known observations in j^{th} treatment

$$\text{Correction factor } (CF) = \frac{(G')^2}{n} = \frac{(y'_{oo} + x)^2}{bv}$$

$$TSS = \sum_{i=1}^{b}\sum_{j=1}^{v} y_{ij}^2 - CF$$

$$= (x^2 + \text{ terms which are constant with respect to } x) \; - \; CF$$

$$SSBl = \frac{1}{v}[(y'_{io} + x)^2 + \text{ terms which are constant with respect to } x] \; - \; CF$$

$$SSTr = \frac{1}{b}[(y'_{oj} + x)^2 + + \text{ terms which are constant with respect to } x] \; - \; CF$$

$$SSE = TSS - SSBl - SSTr$$

$$= x^2 - \frac{1}{v}(y'_{io} + x)^2 - \frac{1}{b}(y'_{oj} + x)^2 + \frac{(y'_{oo} + x)^2}{bv} + \text{ (terms which are constant with respect to } x) \; - \; CF.$$

Find x such that SSE is minimum

$$\frac{\partial(SSE)}{\partial x} = 0 \Rightarrow 2x - \frac{2(y_{io}^{'} + x)}{v} - \frac{2(y_{oj}^{'} + x)}{b} + \frac{2(y_{oo}^{'} + x)}{bv} = 0$$

$$\text{or} \quad x = \frac{vy_{oj}^{'} + by_{io}^{'} - y_{oo}^{'}}{(b-1)(v-1)}$$

The second-order derivative condition for x to provide minimum SSE can be easily verified.

Two missing observations:

If there are two missing observation, then let they be x and y.

- Let the corresponding row sums (block totals) are $(R_1 + x)$ and $(R_2 + y)$.
- Column sums (treatment totals) are $(C_1 + x)$ and $(C_2 + y)$.
- Total of known observations is S.

Then

$$SSE = x^2 + y^2 - \frac{1}{v}[(R_1 + x)^2 + (R_2 + y)^2] - \frac{1}{b}[(C_1 + x)^2 + (C_2 + y)^2] + \frac{1}{bv}(S + x + y)^2$$
$$+ \text{ terms independent of } x \text{ and } y.$$

Now differentiate SSE with respect to x and y, as

$$\frac{\partial(SSE)}{\partial x} = 0 \Rightarrow x - \frac{R_1 + x}{v} - \frac{C_1 + x}{b} + \frac{S + x + y}{bv} = 0$$
$$\frac{\partial(SSE)}{\partial y} = 0 \Rightarrow y - \frac{R_2 + y}{v} - \frac{C_2 + y}{b} + \frac{S + x + y}{bv} = 0.$$

Thus solving the following two linear equations in x and y, we obtain the estimated missing values

$$(b-1)(v-1)x = bR_1 + vC_1 - S - y$$
$$(b-1)(v-1)y = bR_2 + vC_2 - S - x.$$

Adjustments to be done in analysis of variance

(i) Obtain the within block sum of squares from incomplete data.

(ii) Subtract correct error sum of squares from (i). This gives the correct treatment sum of squares.

(iii) Reduce the degrees of freedom of error sum of squares by the number of missing observations.

(iv) No adjustments in other sums of squares are required.

Missing observations in LSD

Let

- x be the missing observation in $(i, j, k)^{th}$ cell, i.e.,

  $y_{ijk}, i = 1,2,..,v,\ j = 1,2,..,v, k = 1,2,..,v.$
- R: Total of known observations in i^{th} row
- C: Total of known observations in j^{th} column
- T: Total of known observation receiving the k^{th} treatment.
- S: Total of known observations

Now

$$\text{Correction factor } (CF) = \frac{(S+x)^2}{v^2}$$

$$\text{Total sum of squares } (TSS) = x^2 + \text{ term which are constant with respect to } x \text{ - } CF$$

$$\text{Row sum of squares } (SSR) = \frac{(R+x)^2}{v} + \text{ term which are constant with respect to } x \text{ - } CF$$

$$\text{Column sum of squares } (SSC) = \frac{(C+x)^2}{v} + \text{ term which are constant with respect to } x \text{ - } CF$$

$$\text{Treatment sum of squares}(SSTr) = \frac{(T+x)^2}{v} + \text{ term which are constant with respect to } x \text{ - } CF$$

$$\text{Sum of squares due to error } (SSE) = TSS - SSR - SSC - SSTr$$

$$= x^2 - \frac{1}{v}\left[(R+x)^2 + (C+x)^2 + (T+x)^2\right] + \frac{2(S+x)^2}{v^2}$$

Choose x such that SSE is minimum. So

$$\frac{d(SSE)}{dx} = 0$$

$$\Rightarrow 2x - \frac{2}{v}\left(R + C + T + 3x\right)) + \frac{4(S+x)}{v^2} = 0$$

$$\text{or } x = \frac{V(R+C+T) - 2S}{(v-1)(v-2)}$$

Adjustment to be done in analysis of variance:

Do all the steps as in the case of RBD.

To get the correct treatment sum of squares, proceed as follows:

- Ignore the treatment classification and consider only row and column classification.
- Substitute the estimated values at the place of missing observation.
- Obtain the error sum of squares from complete data, say SSE_1.
- Let SSE_2 be the error sum of squares based on LSD obtained earlier.
- Find the corrected treatment sum of squares = $SSE_2 - SSE_1$.
- Reduce of degrees of freedom of error sum of squares by the number of missing values.

Chapter 8
Factorial Experiments

Factorial experiments involve simultaneously more than one factor and each factor is at two or more levels. Several factors affect simultaneously the characteristic under study in factorial experiments and the experimenter is interested in the main effects and the interaction effects among different factors.

First, we consider an example to understand the utility of factorial experiments.

Example: Suppose the yield from different plots in an agricultural experiment depends upon

1. (i) variety of crop and
   (ii) type of fertilizer.

Both the factors are in the control of the experimenter.

2. (iii) Soil fertility. This factor is not in the control of the experimenter.

In order to compare different crop varieties

- assign it to different plots keeping other factors like irrigation, fertilizer, etc. fixed and the same for all the plots.
- The conclusions for this will be valid only for the crops grown under similar conditions with respect to the factors like fertilizer, irrigation etc.

In order to compare different fertilizers (or different dosage of fertilizers)

- sow single crop on all the plots and vary the quantity of fertilizer from plot to plot.
- The conclusions will become invalid if different varieties of the crop are sown.
- It is quite possible that one variety may respond differently than another to a particular type of fertilizer.

Suppose we wish to compare

- two crop varieties – a and b, keeping the fertilizer fixed and
- three varieties of fertilizers – A, B and C.

This can be accomplished with two randomized block designs (*RBD*) by assigning the treatments at random to three plots in any block and two crop varieties at random.

The possible arrangement of the treatments may appear can be as follows.

bB	*bA*	*bC*
bC	*bB*	*bA*
bA	*bC*	*bB*

and

aA	*aB*	*aC*
aC	*aA*	*aB*
aB	*aC*	*aA*

With these two *RBDs*,

- the difference among two fertilizers can be estimated
- but the difference among the crop varieties cannot be estimated. The difference among the crop varieties is entangled with the difference in blocks.

On the other hand, if we use three sets of three blocks each and each block having two plots, then

- randomize the varieties inside each block and
- assign treatments at random to three sets.

The possible arrangement of treatment combinations in blocks can be as follows:

bB	*aB*
aB	*bB*
bB	*aB*

,

aC	*bC*
bC	*aC*
aC	*bC*

and

aA	*bA*
bA	*aA*
bA	*aA*

Here the difference between crop varieties is estimable but the difference between fertilizer treatment is not estimable.

Factorial experiments overcome this difficulty and combine each crop with each fertilizer treatment. There are six treatment combinations as

aA, aB, aC, bA, bB, bC.

Keeping the total number of observations to be 18 (as earlier), we can use *RBD* with three blocks with six plots each, e.g.

bA	*aC*	*aB*	*bB*	*aA*	*bC*
aA	*aC*	*bC*	*aB*	*bB*	*bA*
bB	*aB*	*bA*	*aC*	*aA*	*bC*

Now we can estimate

- the difference between crop varieties and
- the difference between fertilizer treatments. Factorial experiments involve simultaneously more than one factor each at two or more levels.

If the number of levels for each factor is the same, we call it is a **symmetrical factorial experiment.**
If the number of levels of each factor is not the same, then we call it as a **symmetrical or mixed factorial experiment**.

We consider only symmetrical factorial experiments.

Through the factorial experiments, we can study

- the individual effect of each factor and
- interaction effect.

Now we consider a 2^2 factorial experiment with an example and try to develop and understand the theory and notations through this example.

General notation for representing the factors is to use capital letters, e.g., *A, B, C* etc. and levels of a factor are represented in small letters. For example, if there are two levels of *A,* they are denoted as a_0 and a_1. Similarly, the two levels of B are represented as b_0 and b_1**.** Another alternative representation to indicate the two levels of *A* is 0 (for a_0) and 1 (for a_1). The factors of B are then 0 (for b_0) and 1 (for b_1).

Note: An important point to remember is that the factorial experiments are conducted in the design of an experiment. For example, the factorial experiment is conducted as an RBD.

Factorial experiments with factors at two levels (2^2 factorial experiment):

Suppose in an experiment, the values of current and voltage in an experiment affect the rotation per minutes (rpm) of fan speed. Suppose there are two levels of current.

- 5 Ampere, call it as level 1 (C_1) and denote it as a_0
- 10 Ampere, call it as level 2 (C_1) and denote it as a_1.

Similarly, the two levels of voltage are

- 200 volts, call it as level 1 (V_0) and denote it as b_0
- 220 volts, call it as level 2 (V_1) and denote it as b_1.

The two factors are denoted as A, say for current and B, say for voltage.

In order to make an experiment, there are 4 different combinations of values of current and voltage.

1. Current = 5 Ampere and Voltage = 200 Volts, denoted as $C_0V_0 \equiv a_0b_0$
2. Current = 5 Ampere and Voltage = 220 Volts, denoted as $C_0V_1 \equiv a_0b_1$.
3. Current = 10 Ampere and Voltage = 200 Volts, denoted as $C_1V_0 \equiv a_1b_0$
4. Current = 10 Ampere and Voltage = 220 Volts, denoted as $C_1V_1 \equiv a_1b_1$

The responses from those treatment combinations are represented by $a_0b_0 \equiv (1), (a_0b_1) \equiv (b)$, $(a_1b_0) \equiv (a)$ and $(a_1b_1) \equiv (ab)$, respectively.

Now consider the following:

I. $\frac{(C_oV_o)+(C_oV_1)}{2}$: Average effect of voltage for the current level C_0

$$: \frac{(a_ob_o)+(a_ob_1)}{2} \equiv \frac{(1)+(b)}{2}$$

II. $\frac{(C_1V_o)+(C_1V_1)}{2}$: Average effect of voltage for the current level C_1

$$: \frac{(a_1b_o)+(a_1b_1)}{2} \equiv \frac{(a)+(ab)}{2}$$

Compare these two group means (or totals) as follows:

Average effect of V_1 level – Average effect at V_0 level

$$= \frac{(b)+(ab)}{2} - \frac{(1)+(a)}{2}$$
$$= \text{Main effect of voltage}$$
$$= \text{Main effect of } B.$$

Comparison like

$(C_oV_1) - (C_oV_o) \equiv (a) - (1)$: indicate the effect of voltage at current level C_o

and

$(C_1V_1) - (C_1V_0) \equiv (ab) - (b)$: indicate the effect of voltage at current level C_1.

The average interaction effect of voltage and current can be obtained as

$$\begin{pmatrix}\text{Average effect of voltage}\\ \text{at current level } I_o\end{pmatrix}-\begin{pmatrix}\text{Average effect of voltage}\\ \text{at current level } I_1\end{pmatrix}$$

$$= \text{Average effect of voltage at different levels of current.}$$

$$=\frac{(C_1V_1)-(C_1V_o)}{2}-\frac{(C_oV_1)-(C_oV_o)}{2}$$

$$=\frac{(ab)-(b)}{2}-\frac{(a)-(1)}{2}$$

$$=\text{Average interaction effect.}$$

Similarly

$$\frac{(C_oV_o)+(C_1V_o)}{2}=\frac{(1)+(b)}{2}:\text{Average effect of current at voltage level } V_0.$$

$$\frac{(C_oV_1)+(C_1V_1)}{2}=\frac{(a)+(ab)}{2}:\text{Average effect of current at voltage level } V_1$$

Comparison of these two as

$$\begin{pmatrix}\text{Average effect of current}\\ \text{at voltage level } V_0\end{pmatrix}-\begin{pmatrix}\text{Average effect of current}\\ \text{at voltage level } V_1\end{pmatrix}$$

$$=\frac{(C_0V_1)+(C_1V_1)}{2}-\frac{(C_0V_0)+(C_1V_0)}{2}$$

$$=\frac{(a)+(ab)}{2}-\frac{(1)+(b)}{2}$$

$$=\text{Main effect of current}$$

$$=\text{Main effect of } A.$$

Comparison like

$$(C_1V_0)-(C_0V_0)=(b)-(1):\text{Effect of current at voltage level } V_0$$

$$(C_1V_1)-(C_0V_1)=(ab)-(a):\text{Effect of current at voltage level } V_1$$

:

The average interact effect of current and voltage can be obtained as

$$\begin{pmatrix}\text{Average effect of current}\\ \text{at voltage level } V_0\end{pmatrix}-\begin{pmatrix}\text{Average effect of current}\\ \text{at voltage level } V_1\end{pmatrix}$$

$$=\text{Average effect of current at different levels of voltage}$$

$$=\frac{(C_1V_1)-(C_0V_1)}{2}-\frac{(C_1V_0)-(C_0V_0)}{2}$$

$$=\frac{(ab)-(a)}{2}-\frac{(b)-(1)}{2}$$

$$=\text{Average interaction effect}$$

$$=\text{Same as average effects of voltage at different levels of current.}$$

(It is expected *that* the interaction effect of current and voltage is same as the interaction effect of voltage and current).

The quantity

$$\frac{(C_0V_0)+(C_1V_0)+(C_0V_1)+(C_1V_1)}{4}=\frac{(1)+(a)+(b)+(ab)}{4}$$

gives the **general mean effect** of all the treatment combination.

Treating (ab) as $(a)(b)$ symbolically (mathematically and conceptually, it is incorrect), we can now express all the main effects, interaction effect and general mean effect as follows:

$$\text{Main effect of } A=\frac{(a)+(ab)}{2}-\frac{(1)+(b)}{2}=\frac{1}{2}\left[(ab)-(b)+(a)-(1)\right]=\frac{(a-1)(b+1)}{2}$$

$$\text{Main effect of } B=\frac{(b)+(ab)}{2}-\frac{(1)+(a)}{2}=\frac{1}{2}\left[(ab)-(a)+(b)-(1)\right]=\frac{(a+1)(b-1)}{2}$$

$$\text{Interaction effect of } A \text{ and } B=\frac{(ab)-(b)}{2}-\frac{(a)-(1)}{2}=\frac{1}{2}\left[(ab)-(a)+(1)-(b)\right]=\frac{(a-1)(b-1)}{2}$$

$$\text{General mean effect } (M)=\frac{(1)+(a)+(b)+(ab)}{4}=\frac{1}{4}\left[(1)+(a)+(b)+(ab)\right]=\frac{(a+1)(b+1)}{4}$$

Notice the roles of + and – signs as well as the divisor.

- There are two effects related to *A* and *B*.
- To obtain the effect of a factor, write the corresponding factor with – sign and others with + sign. For example, in the main effect of *A*, *a* occurs with – sign as in $(a - 1)$ and *b* occurs with + sign as in $(b + 1)$.
- In *AB*, both the effects are present so *a* and *b* both occur with + signs as in $(a + 1)(b + 1)$.
- Also note that the main and interaction effects are obtained by considering the typical differences of averages, so they have divisor 2 whereas the general mean effect is based on all the treatment combinations and so it has divisor 4.
- There is a well defined statistical theory behind this logic but this logic helps in writing the final treatment combination easily. This is demonstrated later with appropriate reasoning.

Other popular notations of treatment combinations are as follows:

$a_0b_0 \equiv 0\ \ 0 \equiv I$
$a_0b_1 \equiv 0\ \ 1 \equiv a$
$a_1b_0 \equiv 1\ \ 0 \equiv b$
$a_1b_1 \equiv 1\ \ 1 \equiv ab.$

Sometimes 0 is referred to as ‘low level’ and 1 is referred to ‘high level’.

Here *I* denote that both factors are at lower levels (a_0b_0 or 0 0). This is called as the **control treatment.**

These effects can be represented in the following table

Factorial effects	Treatment combinations				Divisor
	(1)	(a)	(b)	(ab)	
M	+	+	+	+	4
A	-	+	-	+	2
B	-	-	+	+	2
AB	+	-	-	+	2

The model corresponding to 2^2 factorial experiment is

$$y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + \varepsilon_{ijk},\ i = 1,2,\ j = 1,2,\ k = 1,2,...,n$$

where n observations are obtained for each treatment combinations.

When the experiments are conducted factor by factor, then much more resources are required in comparison to the factorial experiment. For example, if we conduct *RBD* for three-levels of voltage V_0, V_1 and V_2 and two levels of current I_0 and I_1, then to have 10 degrees of freedom for the error variance, we need

- 6 replications on voltage
- 11 replications on current.

So the total number of fans needed is 40.

For the factorial experiment with 6 combinations of 2 factors, the total number of fans needed are 18 for the same precision.

We have considered the situation up to now by assuming only one observation for each treatment combination, i.e., no replication. If r replicated observations for each of the treatment combinations are obtained, then the expressions for the main and interaction effects can be expressed as

$$A = \frac{1}{2r}\left[(ab) + (a) - b - (1)\right]$$

$$B = \frac{1}{2r}\left[(ab) + (b) - a - (1)\right]$$

$$AB = \frac{1}{2r}\left[(ab) + (1) - a - (b)\right]$$

$$M = \frac{1}{4r}\left[(ab) + (a) + (b) + (1)\right].$$

Now we detail the statistical theory and concepts related to these expressions.

Let $Y_* = ((1),\ a,\ b,\ ab)'$ be the vector of total response values. Then

$$A = \frac{1}{2r}\ell_A' Y_* = \frac{1}{2r}(-1 \quad 1 \quad -1 \quad 1)Y_*$$
$$B = \frac{1}{2r}\ell_B' Y_* = \frac{1}{2r}(-1 \quad -1 \quad 1 \quad 1)Y_*$$
$$AB = \frac{1}{2r}\ell_{AB}' Y_* = \frac{1}{2r}(1 \quad -1 \quad -1 \quad 1)Y_*.$$

Note that A, B and AB are the linear contrasts. Recall that a linear parametric function is estimable only when it is in the form of linear contrast. Moreover, A, B and AB are the linear orthogonal contrasts in the total response values $(1),\ a,\ b,\ ab$ except for the factor $1/2r$.

The sum of squares of a linear parametric function $\ell' y$ is given by $\frac{(\ell' y)^2}{\ell'\ell}$. If there are r replicates, then the sum of squares is $\frac{(\ell' y)^2}{r\ell'\ell}$. It may also be recalled under the normality of y's, this sum of squares has a Chi-square distribution with one degree of freedom (χ_1^2). Thus the various associated sum of squares due to A, B and AB are given by the following:

$$SSA = \frac{(\ell_A' Y_*)^2}{r\ell_A'\ell_A} = \frac{1}{4r}(ab + a - b - (1))^2$$
$$SSB = \frac{(\ell_B' Y_*)^2}{r\ell_B'\ell_B} = \frac{1}{4r}(ab + b - a - (1))^2$$
$$SSAB = \frac{(\ell_{AB}' Y_*)^2}{r\ell_{AB}'\ell_{AB}} = \frac{1}{4r}(ab + (1) - a - b)^2.$$

Each of SSA, SSB and $SSAB$ has χ_1^2 under normality of Y_*.

The sum of squares due to total is computed as usual

$$TSS = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{r} y_{ijk}^2 - \frac{G^2}{4r}$$

where

$$G = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{r} y_{ijk}$$

is the grand total of all the observations.

The TSS has χ^2 distribution with $(2^2 r - 1)$ degrees of freedom.

The sum of squares due to error is also computed as usual as

$$SSE = TSS - SSA - SSB - SSAB$$

which has χ^2 distribution with

$$(4r-1)-1-1-1 = 4(r-1)$$

degrees of freedom.

The mean squares are

$$MSA = \frac{SSA}{1},$$

$$MSB = \frac{SSB}{1},$$

$$MSAB = \frac{SSAB}{1},$$

$$MSE = \frac{SSA}{4(r-1)}.$$

The F -statistic corresponding to A, B and AB are

$$F_A = \frac{MSA}{MSE} \sim F(1, 4(r-1) \text{ under } H_0,$$

$$F_B = \frac{MSB}{MSE} \sim F(1, 4(r-1) \text{ under } H_0,$$

$$F_{AB} = \frac{MSAB}{MSE} \sim F(1, 4(r-1) \text{ under } H_0.$$

The ANOVA table is case of 2^2 factorial experiment is given as follows:

Source	Sum of squares	Degrees of freedom	Mean squares	F
A	SSA	1	MSA	$F_A = \frac{MSA}{MSE}$
B	SSB	1	MSB	$F_B = \frac{MSB}{MSE}$
AB	$SSAB$	1	$MSAB$	$F_{AB} = \frac{MSAB}{MSE}$
$Error$	SSE	$4(r-1)$	MSE	
Total	TSS	$4r-1$		

The decision rule is to reject the concerned null hypothesis when the value of the concerned F statistic

$$F_{\text{effect}} > F_{1-\alpha}(1, 4(r-1)).$$

2^3 Factorial experiment:

Suppose that in a complete factorial experiment, there are three factors - A, B and C, each at two levels, viz., $a_0, a_1; b_0, b_1$ and c_0, c_1 respectively. There are a total of eight number of combinations:

$a_0b_0c_0$, $a_0b_0c_1$, $a_0b_1c_0$, $a_0b_1c_1$,
$a_1b_0c_0$, $a_1b_0c_1$, $a_1b_1c_0$, $a_1b_1c_1$.

Each treatment combination has r replicates, so the total number of observations are $N = 2^3 r = 8r$ that are to be analyzed for their influence on the response.

Assume the total response values are

$$Y_* = \left[(1),\ a,\ b,\ ab,\ c,\ ac,\ bc,\ abc\right]'.$$

The response values can be arranged in a three-dimensional contingency table. The effects are determined by the linear contrasts

$$\ell'_{effect} Y_* = \ell'_{effect}\left((1),\ a,\ b,\ ab,\ c,\ ac,\ bc,\ abc\right)$$

using the following table:

Factorial effect	Treatment combinations							
	(1)	*a*	*b*	*ab*	*c*	*ac*	*bc*	*abc*
I	+	+	+	+	+	+	+	+
A	-	+	-	+	-	+	-	+
B	-	-	+	+	-	-	+	+
AB	+	-	-	+	+	-	-	+
C	-	-	-	-	+	+	+	+
AC	+	-	+	-	-	+	-	+
BC	+	+	-	-	-	-	+	+
ABC	-	+	+	-	+	-	-	+

Note that once a few rows have been determined in this table, rest can be obtained by simple multiplication of the symbols. For example, consider the column corresponding to *a*, we note that A has + sign, B has – sign , so AB has – sign (sign of $A\times$ sign of B). Once AB has - sign, C has – sign then ABC has (sign of AB X sign of C) which is + sign and so on.

The first row is a basic element. With this $a = 1'Y_*$ can be computed where 1 is a column vector of all elements unity. If other rows are multiplied with the first row, they stay unchanged (therefore we call it as identity and denoted as I). Every other row has the same number of + and – signs. If + is replaced by 1 and – is replaced by -1, we obtain the vectors of orthogonal contrasts with the norm $8(= 2^3)$.

If each row is multiplied by itself, we obtain I (first row). The product of any two rows leads to a different row in the table. For example

$$A.B = AB$$
$$AB.B = AB^2 = A$$
$$AC.BC = A.C^2BB = AB.$$

The structure in the table helps in estimating the average effect.

For example, the average effect of A is

$$A = \frac{1}{4r}\left[(a) - (1) + (ab) - (b) + (ac) - (c) + (abc) - (bc)\right]$$

which has the following explanation.

(i) Average effect of A at low level of B and $C \equiv (a_1b_0c_0) - (a_0b_0c_0) \equiv \frac{\left[(a) - (1)\right]}{r}$.

(ii) Average effect of A at high level of B and low level of $C \equiv (a_1b_1c_0) - (a_0b_1c_0) \equiv \frac{\left[(ab) - (b)\right]}{r}$

(iii) Average effect of A at low level of B and high level of $C \equiv (a_1b_0c_1) - (a_0b_0c_1) \equiv \frac{\left[(ac) - (c)\right]}{r}$.

(iv) Average effect of A at high level of B and $C \equiv (a_1b_1c_1) - (a_0b_1c_1) \equiv \frac{\left[(abc) - (bc)\right]}{r}$.

Hence for all combinations of B and C, the average effect of A is the average of all the average effects in (i)-(iv).

Similarly, other main and interaction effects are as follows:

$$B = \frac{1}{4r}\left[(b) + (ab) + (bc) + (abc) - (1) - (a) - (c) - (ac)\right] = \frac{(a+1)(b-1)(c+1)}{4r}$$
$$C = \frac{1}{4r}\left[c + (ac) + (bc) + (abc) - (1) - (a) - (b) - (ab)\right] = \frac{(a+1)(b+1)(c-1)}{4r}$$
$$AB = \frac{1}{4r}\left[(1) + (ab) + (c) + (abc) - (a) - (b) - (ac) - (bc)\right] = \frac{(a-1)(b-1)(c+1)}{4r}$$
$$AC = \frac{1}{4r}\left[(1) + (b) + (ac) + (abc) - (a) - (ab) - (c) - (bc)\right] = \frac{(a-1)(b+1)(c-1)}{4r}$$
$$BC = \frac{1}{4r}\left[(1) + (a) + (bc) + (abc) - (b) - (ab) - (c) - (ac)\right] = \frac{(a+1)(b-1)(c-1)}{4r}$$
$$ABC = \frac{1}{4r}\left[(abc) + (ab) + (b) + (c) - (ab) - (ac) - (bc) - (1)\right] = \frac{(a-1)(b-1)(c-1)}{4r}.$$

Various sum of squares in the 2^3 factorial experiment are obtained as

$$SS(\text{Effect}) = \frac{(\text{linear contrast})^2}{8r} = \frac{\left(\ell'_{\text{effect}} Y_*\right)^2}{r\ell'_{\text{effect}}\ell_{\text{effect}}}$$

which follow a Chi-square distribution with one degree of freedom under normality of Y_*. The corresponding mean sum of squares is obtained as

$$MS(\text{Effect}) = \frac{SS(\text{Effect})}{\text{Degrees of freedom}}.$$

The corresponding F -statistics are obtained by

$$F_{\text{Effect}} = \frac{MS(\text{Effect})}{MS(\text{Error})}$$

which follows an $F-$distribution with degrees of freedoms 1 and error degrees of freedom under the respective null hypothesis. The decision rule is to reject the corresponding null hypothesis at the α level of significance whenever

$$F_{effect} > F_{1-\alpha}(1, df_{\text{error}}).$$

These outcomes are presented in the following ANOVA table

Sources	Sum of squares	Degrees of freedom	Mean sum of squares	F
A	SSA	1	$MSA = SSA/1$	F_A
B	SSB	1	$MSB = SSB/1$	F_B
AB	$SSAB$	1	$MSAB = SSAB/1$	F_{AB}
C	SSC	1	$MSC = SSC/1$	F_C
AC	$SSAC$	1	$MSAC = SSAC/1$	F_{AC}
BC	$SSBC$	1	$MSBC = SSBC/1$	F_{BC}
ABC	$SSABC$	1	$MSABC = SSABC/1$	F_{ABC}
Error	$SS(Error)$	$8(r-1)$	$MS(Error) = SS(Error)/\{8(r-1)\}$	
Total	TSS	$8r-1$		

2^n Factorial experiment:

Based on the theory developed for 2^2 and 2^3 factorial experiments, we now extend them for the 2^n factorial experiment.

- Capital letters $A, B, C, \ldots$ denote the **factors**. They are the main effect contrast for the factors $A, B, C, \ldots$
- $AB, AC, BC, \ldots$ denote the first order or **2-factor interactions**
- $ABC, ABD, BCD, \ldots$ denote the **second-order** or **3-factor interactions** and so on.
- Each of the main effect and interaction effect carries one degree of freedom.
- Total number of main effects = $\binom{n}{1} = n.$
- Total number of first-order interactions = $\binom{n}{2}$.
- Total number of second-order interactions = $\binom{n}{3}$

and so on.

Standard order for treatment combinations:

The list of treatments can be expressed in a standard order.

- For one factor A, the standard order is $(1),\ a$.
- For two factors *A* and *B,* the standard order is obtained by adding *b* and *ab* in the standard order of one factor *A*. This is derived by multiplying (1) and *a* by *b* , i.e.
  $b \times \{(1), a\} = (1),\ a,\ b,\ ab$.
- For three factors, add $c,\ ac, bc$ and abc which are derived by multiplying the standard order of *A* and *B* by *c*, i.e.
  $c \times \{(1,\ a,\ b,\ ab\} = (1),\ a,\ b,\ ab,\ c,\ ac,\ bc,\ abc$.

Thus the standard order of any factor is obtained step by step by multiplying it with an additional letter to preceding standard order.

For example, the standard order of *A, B, C* and *D* is 2^4 factorial experiment is

$$(1),\ a,\ b,\ ab,\ c,\ ac,\ bc,\ abc,\ d \times \{(1),\ a,\ b,\ ab,\ c,\ ac,\ bc,\ abc\}$$
$$= (1),\ a,\ b,\ ab,\ c,\ ac,\ bc,\ abc,\ d,\ ad,\ bd,\ abd,\ cd,\ acd,\ bcd,\ abcd.$$

How to find the contrasts for main effects and interaction effect:

Recall that earlier, we had illustrated the concept in writing the contrasts for main and interaction effects. For example, in a 2^2 factorial experiment, we had expressed

$$A = \frac{1}{2}(a-1)(b+1) = \frac{1}{2}\left[(-1)+(a)-(b)+(ab)\right]$$
$$AB = \frac{1}{2}(a-1)(b-1) = \frac{1}{2}\left[(1)-(a)-(b)+(ab)\right].$$

Note that each effect has two component - divisor and contrast. When the order of factorial increases, it is cumbersome to derive such expressions. Some methods have been suggested to write the expressions for factorial effects. First, we detail how to write divisor and then illustrate the methods for obtaining the contrasts.

How to write divisor:

In a 2^n factorial experiment,

- the general mean effect has divisor 2^n and
- any effect (main or interaction) has divisor 2^{n-1}.

For example, in a 2^6 factorial experiment, the general mean effect has divisor 2^6 and any main effect or interaction effect of any order has divisor $2^{6-1} = 2^5$.

If r replicates of each effect are available, then

- the general mean effect has divisor $r2^n$ and
- any main effect or interaction effect of any order has a divisor $r2^{n-1}$.

How to write contrasts:

Method 1:

Contrast belonging to the main effects and the interaction effects are written as follows:

$$A = (a-1)(b+1)(c+1)...(z+1)$$
$$B = (a+1)(b-1)(c+1)...(z+1)$$
$$C = (a+1)(b+1)(c-1)...(z+1)$$
$$\vdots$$
$$AB = (a-1)(b-1)(c+1)...(z+1)$$
$$BC = (a+1)(b-1)(c-1)...(z+1)$$
$$\vdots \qquad\qquad \vdots$$
$$ABC = (a-1)(b-1)(c-1)...(z+1).$$
$$\vdots$$
$$ABC...Z = (a-1)(b-1)(c-1)...(z-1)$$

Look at the pattern of assigning + and – signs on the right-hand side. The letters common on left and right hand sides of the equality (=) sign (irrespective of small or capital letters) contain – sign and rest contain + sign.

The expression on right-hand side when simplified algebraically give the contrasts in terms of treatment combination. For example, in a 2^3 factorial

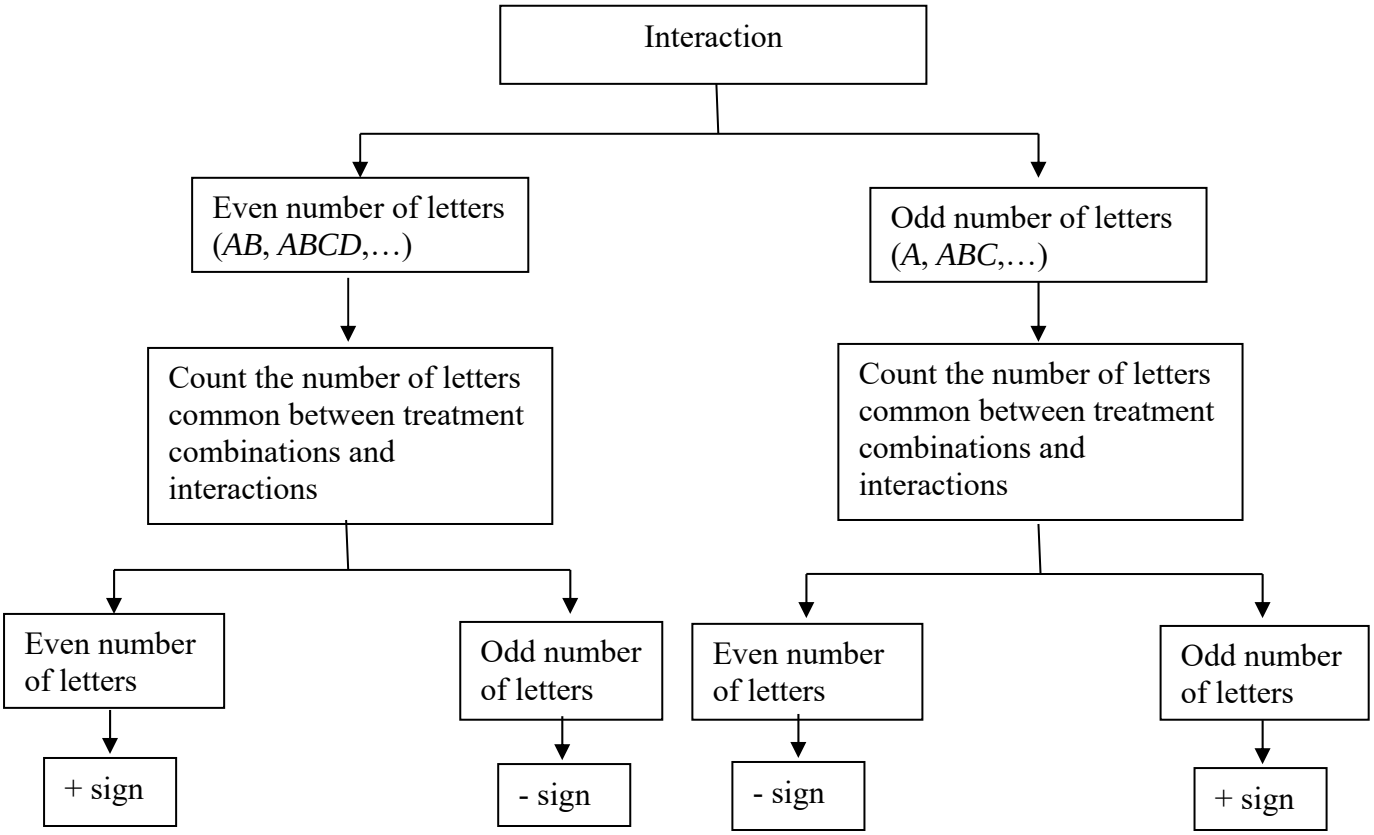
$$A = \frac{1}{2^{3-1}}(a-1)(b+1)(c+1)$$
$$= \frac{1}{4}\left[-(1)+(a)-(b)+(ab)-(c)+(ac)-(bc)+(abc)\right]$$
$$M = \frac{1}{2^3}(a+1)(b+1)(c+1)$$
$$= \frac{1}{8}\left[(1)+(a)+(b)+(ab)+(c)+(ac)+(bc)+(abc)\right]$$

Method 2

- Form a table such that
  - rows correspond to the main or interaction effect and
  - columns correspond to treatment combinations (or another way round)
- + and – signs in the table indicate the sign of the treatment combinations of main and interaction effects.
- Signs are determined by the "rule of odds and evens" given as follows:
  - if the interaction has an even number of letters $(AB, ABCD, ...)$, a treatment combination having an even number of letters common with the interaction enters with a + sign and one with an odd number of letters common enters with a – sign.
  - if the interaction has an odd number of letters $(A, ABC, ...)$, the rule is reversed.
- Once few rows are filled up, others can be obtained through multiplication rule. For example, the sign of $ABCD$ is obtained as

  (sign of $A\times$ sign of BCD) or (sign of $AB\times$ sign of CD).
- Treatment combination (1) is taken to have an even number (zero) of letters common with every interaction.

This rule of assignment of + or – is illustrated in the following flow diagram:



For example, in a 2^3 factorial experiment, write

- rows for main and interaction effects and
- columns for treatment combinations in standard order.
- Take treatment combination (1) to have an even number (zero) of letter common with every interaction.

This gives the following table

Factorial effect	Treatment combinations							
	(1)	*a*	*b*	*ab*	*c*	*ac*	*bc*	*abc*
I	+	+	+	+	+	+	+	+
A	-	+	-	+	-	+	-	+
B	-	-	+	+	-	-	+	+
AB	+	-	-	+	+	-	-	+
C	-	-	-	-	+	+	+	+
AC	+	-	+	-	-	+	-	+
BC	+	+	-	-	-	-	+	+
ABC	-	+	+	-	+	-	-	+

Sums of squares:

Suppose 2^n factorial experiment is carried out in a randomized block design with r replicates.

Denote the total yield (output) from r plots (experimental units) receiving a particular treatment combination by the same symbol within a square bracket. For example, $[ab]$ denotes the total yield from the plots receiving the treatment combination (ab).

In a 2^2 factorial experiment, the factorial effect totals are

$$[A]=[ab]-[b]+[a]-[1]$$

$[ab]$ = treatment total, i.e. the sum of r observations in which both the factors A and B are at the second level.

$[a]$ = treatment total, i.e., the sum of r observations in which factor A is at the second level and factor B is at the first level

$[b]$ = treatment total, i.e., the sum of r observations in which factor A is at the first level and factor B is at the second level.

$[1]$ = treatment total, i.e. the sum of r observations in which both the factors A and B are at the first level.

$$[A]=\sum_{i=1}^{r}\left[y_{i(ab)}-y_{i(b)}+y_{i(a)}-y_{i(1)}\right]$$
$$=\ell_A' y_A \text{ (say).}$$

where ℓ_A is a vector of +1 and =1 and y_A is a vector denoting the responses from *ab, b, a* and 1. Similarly, other effects can also be found.

Similarly, other effects can also be found.

The sum of squares due to a particular effect is obtained as

$$\frac{[\text{Total yield}]^2}{\text{Total number of observations}}.$$

In a 2^2 factorial experiment in an RBD, the sum of squares due to A is

$$SSA=\frac{(\ell_A' y_A)^2}{r2^2}.$$

In a 2^n factorial experiment in an RBD, the divisor will be $r.2^n$. If Latin square design is used based on $2^n \times 2^n$ Latin square, then r is replaced by 2^n.

Yates method of computation of the sum of squares:

Yates method gives a systematic approach to find the sum of squares. We are not presenting here the complete method. Only the part which is used for computing only the sum of squares is presented and the method to verify them is not presented.

It has the following steps

1. First, write the treatment combinations in the standard order in the column at the beginning of the table, called a **treatment column.**

2. Find the total yield for each treatment. Write this as the second column of the table, called a **yield column.**

3. Obtain columns (1),(2),…,(n) successively

   - (i) obtain column (1) from yield column

     (a) upper half is obtained by adding yields in pairs.

     (b) the second half is obtained by taking differences in pairs, the difference obtained by subtracting the first term of pairs from the second term.

   (ii) The columns (2),(3),…,(n) are obtained from preceding ones in the same manner as used for getting (1) from the yield columns.

4. This process of finding columns is repeated n times in 2^n factorial experiment.

5. Sum of squares due to interaction $= \dfrac{\left[\text{column}(n)\right]^2}{\text{Total number of observations}}$

Example: Yates procedure for a 2^2 factorial experiment

Treatment Combination	Yield (total from all r replicates)	(1)	(2)
(1)	(1)	(1) + (a)	(1) + (a) + (b)+ (ab) = [M]
a	(a)	(b) + (ab)	-(1) + (a) - (b) + (ab) = [A]
b	(b)	(a) - (1)	-(1) - (a) + (b) + (ab) = [B]
ab	(ab)	(ab) - (b)	(1) - (a) - (b) + (ab) = [AB]

Note: The columns are repeatedly obtained 2 times due to 2^2 factorial experiment.

Now $SSA = \dfrac{\left[A\right]^2}{4r}$; $SSB = \dfrac{\left[B\right]^2}{4r}$; $SSAB = \dfrac{\left[AB\right]^2}{4r}$

Example: Yates procedure for 2^3 factorial experiment.

Treatment	Yield (total from all r replicates)				
(1)	(2)	(3)	(4)	(5)	(6)
1	(1)	$u_1 = (1) + (a)$	$v_1 = u_1 + u_2$	$w_1 = v_1 + v_2$	$[M]$
a	(a)	$u_2 = (b) + (ab)$	$v_2 = u_3 + u_4$	$w_2 = v_3 + v_4$	$[A]$
b	(b)	$u_3 = (c) + (ac)$	$v_3 = u_5 + u_6$	$w_3 = v_5 + v_6$	$[B]$
ab	(ab)	$u_4 = (bc) + (abc)$	$v_4 = u_7 + u_8$	$w_4 = v_7 + v_8$	$[AB]$
c	(ac)	$u_5 = (a) - (1)$	$v_5 = u_2 - u_7$	$w_5 = v_2 - v_1$	$[C]$
ac	(ac)	$u_6 = (ab) - (b)$	$v_6 = u_4 - u_3$	$w_6 = v_4 - v_3$	$[AC]$
bc	(bc)	$u_7 = (ac) - (c)$	$v_7 = v_6 - u_5$	$w_7 = v_6 - v_5$	$[BC]$
abc	(abc)	$u_8 = (abc) - (bc)$	$v_8 = u_8 - u_7$	$w_8 = v_8 - v_7$	$[ABC]$

The sum of squares are obtained as follows when the design is RBD.

$$SS(Effect) = \frac{[\text{Effect}]^2}{r.2^3}.$$

For analysis of 2^n factorial experiment, the analysis of variance involves the partitioning of treatment sum of squares so as to obtain sum of squares due to main and interaction effects of factors. These sum of squares are mutually orthogonal, so

Treatment SS = Total of SS due to main and interaction effects.

For example:

In 2^2 factorial experiment in an RBD with r replications, the division of degrees of freedom and the treatment sum of squares are as follows:

Source	Degree of freedom	Sum of squares
Replications	$r-1$	
Treatments	$4-1=3$	
- A	1	$[A]^2/4r$
- B	1	$[B]^2/4r$
- AB	**1**	$[AB]^2/4r$
Error	$3(r-1)$	
Total	$4r-1$	

The decision rule is to reject the concerned null hypothesis when the related F-statistic $F_{effect} > F_{1-\alpha}(1, 3(r-1))$.

Chapter 9
Confounding

If the number of factors or levels increase in a factorial experiment, then the number of treatment combinations increases rapidly. When the number of treatment combinations is large, then it may be difficult to get the blocks of sufficiently large size to accommodate all the treatment combinations. Under such situations, one may use either connected incomplete block designs, e.g., balanced incomplete block designs (BIBD) where all the main effects and interaction contrasts can be estimated or use unconnected designs where not all these contrasts can be estimated.

Non-estimable contrasts are said to be confounded.

Note that a linear function $\lambda'\beta$ is said to be estimable if there exist a linear function $l'y$ of the observations on random variable y such that $E(l'y) = \lambda'\beta$. Now there arise two questions. Firstly, what does confounding means and secondly, how does it compares to using BIBD.

In order to understand the confounding, let us consider a simple example of 2^2 factorial with factors a and b. The four treatment combinations are (1), a, b and ab. Suppose each batch of raw material to be used in the experiment is enough only for two treatment combinations to be tested. So two batches of raw material are required. Thus two out of four treatment combinations must be assigned to each block. Suppose this 2^2 factorial experiment is being conducted in a randomized block design. Then the corresponding model is

$$E(y_{ij}) = \mu + \beta_i + \tau_j,$$

then

$$A = \frac{1}{2r}\left[ab + a - b - (1)\right],$$

$$B = \frac{1}{2r}\left[ab + b - a - (1)\right],$$

$$AB = \frac{1}{2r}\left[ab + (1) - a - b\right].$$

Suppose the following block arrangement is opted:

Block 1	Block 2
(1) *ab*	*a* *b*

The block effects of blocks 1 and 2 are β_1 and β_2, respectively, then the average responses corresponding to treatment combinations a, b, ab and (1) are

$$E[y(a)] = \mu + \beta_2 + \tau(a),$$
$$E[y(b)] = \mu + \beta_2 + \tau(b),$$
$$E[y(ab)] = \mu + \beta_1 + \tau(ab),$$
$$E[y(1)] = \mu + \beta_1 + \tau(1),$$

respectively. Here $y(a)$, $y(b)$, $y(ab)$, $y(1)$ and $\tau(a)$, $\tau(b)$, $\tau(ab)$, $\tau(1)$ denote the responses and treatments corresponding to a, b, ab and (1), respectively. Ignoring the factor $1/2r$ in A, B, AB and using $E[y(a)], E[y(b)], E[y(ab)], E(y(1)]$, the effect A is expressible as follows :

$$\begin{aligned} A &= [\mu + \beta_1 + \tau(ab)] + [\mu + \beta_2 + \tau(a)] - [\mu + \beta_2 + \tau(b)] - [\mu + \beta_1 + \tau(1)] \\ &= \tau(ab) + \tau(a) - \tau(b) - \tau(1). \end{aligned}$$

So the block effect is not present in A and it is not mixed up with the treatment effects. In this case, we say that the main effect A is not confounded. Similarly, for the main effect B, we have

$$\begin{aligned} B &= [\mu + \beta_1 + \tau(ab)] + [\mu + \beta_2 + \tau(b)] - [\mu + \beta_2 + \tau(a)] - [\mu + \beta_1 + \tau(1)] \\ &= \tau(ab) + \tau(b) - \tau(a) - \tau(1). \end{aligned}$$

So there is no block effect present in B and thus B is not confounded. For the interaction effect AB, we have

$$\begin{aligned} AB &= [\mu + \beta_1 + \tau(ab)] + [\mu + \beta_1 + \tau(1)] - [\mu + \beta_2 + \tau(a)] - [\mu + \beta_2 + \tau(b)] \\ &= 2(\beta_1 - \beta_2) + \tau(ab) + \tau(1) - \tau(a) - \tau(b). \end{aligned}$$

Here the block effects are present in AB. In fact, the block effects are β_1 and β_2 are mixed up with the treatment effects and cannot be separated individually from the treatment effects in AB. So AB is said to be confounded (or mixed up) with the blocks.

Alternatively, if the arrangement of treatments in blocks are as follows:

Block 1	Block 2
ab	(1)
a	b

then the main effect A is expressible as

$$A = [\mu + \beta_1 + \tau(ab)] + [\mu + \beta_1 + \tau(a)] - [\mu + \beta_2 + \tau(b)] - [\mu + \beta_2 + \tau(1)]$$
$$= 2(\beta_1 - \beta_2) + \tau(ab) + \tau(a) - \tau(b) - \tau(1)$$

Observe that the block effects β_1 and β_2 are present in this expression. So the main effect A is confounded with the blocks in this arrangement of treatments.
So the main effect A is confounded with the blocks in this arrangement of treatments.

We notice that it is in our control to decide that which of the effect is to be confounded. The order in which treatments are run in a block is determined randomly. The choice of block to be run first is also randomly decided.

The following observation emerges from the allocation of treatments in blocks:
"*For a given effect, when two treatment combinations with the same signs are assigned to one block and the other two treatment combinations with the same but opposite signs are assigned to another block, then the effect gets confounded".*

For example, in case AB is confounded, then

- ab and (1) with + signs are assigned to block 1 whereas
- a and b with – signs are assigned to block 2.

Similarly, when A is confounded, then

- a and ab with + signs are assigned to block 1 whereas
- (1) and b with – signs are assigned to block 2.

The reason behind this observation is that if every block has treatment combinations in the form of linear contrast, then effects are estimable and thus unconfounded. This is also evident from the theory of linear estimation that a linear parametric function is estimable if it is in the form of a linear contrast.

The contrasts which are not estimable are said to be **confounded with the differences between blocks** (or block effects). The contrasts which are estimable are said to be unconfounded with blocks or free from block effects.

Comparison of balanced Incomplete Block design (BIBD versus factorial:

Now we explain how confounding and BIBD compares together. Consider a 2^3 factorial experiment which needs the block size to be 8. Suppose the raw material available to conduct the experiment is sufficient only for a block of size 4. One can use BIBD in this case with parameters $b=14, k=4, v=8, r=7$ and $\lambda=3$ (such BIBD exists). For this BIBD, the efficiency factor is

$$E=\frac{\lambda v}{kr}=\frac{6}{8}$$

and

$$Var(\hat{\tau}_j-\hat{\tau}_{j'})_{BIBD}=\frac{2k}{\lambda v}\sigma^2=\frac{2}{6}\sigma^2 \quad (j\neq j').$$

Consider now an unconnected design in which 7 out of 14 blocks get treatment combination in block 1 as

$$\boxed{a\ \ b\ \ c\ \ abc}$$

and remaining 7 blocks get treatment combination in block 2 as

$$\boxed{(1)\ \ ab\ \ bc\ \ ac}$$

In this case, all the effects A, B, C, AB, BC and AC are estimable but ABC is not estimable because the treatment combinations with all + and all – signs in

$$\begin{aligned} ABC &= (a-1)(b-1)(c-1) \\ &= \underbrace{(a+b+c+abc)}_{\text{in block 1}} - \underbrace{((1)+ab+bc+ac)}_{\text{in block 2}} \end{aligned}$$

are contained in same blocks. In this case, the variance of estimates of unconfounded main effects and interactions is $8\sigma^2/7$. Note that in case of RBD,

$$Var(\hat{\tau}_j-\hat{\tau}_{j'})_{RBD}=\frac{2\sigma^2}{r}=\frac{2\sigma^2}{7} \quad (j\neq j')$$

and there are four linear contrasts, so the total variance is $4\times(2\sigma^2/7)$ which gives the factor $8\sigma^2/7$ and which is smaller than the variance under BIBD.

We observe that at the cost of not being able to estimate ABC, we have better estimates of A, B, C, AB, BC and AC with the same number of replicates as in BIBD. Since higher order interactions are difficult to interpret and are usually not large, so it is much better to use confounding arrangements which provide better estimates of the interactions in which we are more interested.

Note that this example is for understanding only. As such the concepts behind incomplete block design and confounding are different.

Confounding arrangement:

The arrangement of treatment combinations in different blocks, whereby some pre-determined effect (either main or interaction) contrasts are confounded is called a confounding arrangement.

For example, when the interaction ABC is confounded in a 2^3 factorial experiment, then the confounding arrangement consists of dividing the eight treatment combinations into following two sets:

$\boxed{a \;\; b \;\; c \;\; abc}$

and

$\boxed{(1) \;\; ab \;\; bc \;\; ac}$

With the treatments of each set being assigned to the same block and each of these sets being replicated same number of times in the experiment, we say that we have a confounding arrangement of a 2^3 factorial in two blocks. It may be noted that any confounding arrangement has to be such that only predetermined interactions are confounded and the estimates of interactions which are not confounded are orthogonal whenever the interactions are orthogonal.

Defining contrast:

The interactions which are confounded are called the defining contrasts of the confounding arrangement.

A confounded contrast will have treatment combinations with the same signs in each block of the confounding arrangement. For example, if effect $AB = (a-1)(b-1)(c+1)$ is to be confounded, then put all factor combinations with + sign, i.e., (1), ab, c and abc in one block and all other factor combinations with – sign, i.e., a, b, ac and bc in another block. So the **block size reduces to 4 from 8** when one effect is confounded in 2^3 factorial experiment.

Suppose if along with ABC confounded, we want to confound C also,. To obtain such blocks, consider the blocks where ABC is confounded and divide them into further halves. So the block

$\boxed{a \;\; b \;\; c \;\; abc}$

is divided into following two blocks: $\boxed{a \;\; b}$ and $\boxed{c \;\; abc}$

and the block

$$\boxed{(1)\ \ ab\ \ bc\ \ ac}$$

is divided into following two blocks: $\boxed{(1)\ \ ab}$ and $\boxed{bc\ \ ac}$

These blocks of 4 treatments are divided into 2 blocks with each having 2 treatments and they are obtained in the following way. If only C is confounded then the block with + sign of treatment combinations in C is

$$\boxed{c\ \ ac\ \ bc\ \ abc}$$

and block with – sign of treatment combinations in C is

$$\boxed{(1)\ \ a\ \ b\ \ ab}.$$

Now look into the

(i) following block with + sign when $ABC = (a-1)(b-1)(c-1)$ is confounded,

$$\boxed{a\ \ b\ \ c\ \ abc}$$

(ii) following block with + sign when $C = (a+1)(b+1)(c-1)$ is confounded and

$$\boxed{c\ \ ab\ \ ac\ \ abc}$$

(iii) table of + and – signs in case of 2^3 factorial experiment.

Identify the treatment combinations having common - signs in these two blocks in (i) and (ii). These treatment combinations are are c and abc. So assign them into one block. The remaining treatment combinations out of a, b, c and abc are a and b which go into another block.

Similarly look into the

(a) following block with – sign when ABC is confounded,

$$\boxed{(1)\ \ ab\ \ bc\ \ ac}$$

(b) following block with – sign when C is confounded and

$$\boxed{(1)\ \ a\ \ b\ \ ab}$$

(c) table of + and – signs in case of 2^3 factorial experiment.

Identify the treatment combinations having common – sign in these two blocks in (a) and (b). These treatment combinations are (1) and ab which go into one block and the remaining two treatment combinations ac and bc out of c, ac, bc and abc go into another block. So the blocks where both ABC and C are confounded together are

$\boxed{(1)\ \ ab}$, $\boxed{a\ \ b}$, $\boxed{ac\ \ bc}$ and $\boxed{c\ \ abc}$.

While making these assignments of treatment combinations into four blocks, each of size two, we notice that another effect, viz., AB also gets confounded automatically. Thus we see that when we confound two factors, a third factor is automatically getting confounded. This situation is quite general. The defining contrasts for a confounding arrangement cannot be chosen arbitrarily. If some defining contrasts are selected then some other will also get confounded.

Now we present some definitions which are useful in describing the confounding arrangements.

Generalized interaction:

Given any two interactions, the generalized interaction is obtained by multiplying the factors (in capital letters) and ignoring all the terms with an even exponent.

For example, the generalized interaction of the factors ABC and BCD is $ABC \times BCD = AB^2C^2D = AD$ and the generalized interaction of the factors AB, BC and ABC is $AB \times BC \times ABC = A^2B^3C^2 = B.$

Independent set :

A set of main effects and interaction contrasts is called independent if no member of the set can be obtained as a generalized interaction of the other members of the set.

For example, the set of factors AB, BC and AD is an independent set but the set of factors AB, BC, CD and AD is not an independent set because $AB \times BC \times CD = AB^2C^2D = AD$ which is already contained in the set.

Orthogonal treatment combinations:

The treatment combination $a^p b^q c^r \ldots$ is said to be orthogonal to the interaction $A^x B^y C^z \ldots.$ if $(px + qy + rz + \ldots.)$ is divisible by 2. Since $p, q, r, \ldots, x, y, z, \ldots$ are either 0 or 1, so a treatment combination is orthogonal to an interaction if they have an even number of letters in common.Treatment combination (1) is orthogonal to every interaction.

If $a^{p_1} b^{q_1} c^{r_1} \ldots.$and $a^{p_2} b^{q_2} c^{r_{12}} \ldots$ are both orthogonal to $A^x B^y C^z \ldots,$ then the product $a^{p_1+p_2} b^{q_1+q_2} c^{r_1+r_2} \ldots$ is also orthogonal to $A^x B^y C^z \ldots$ Similarly, if two interactions are orthogonal to a treatment combination, then their generalized interaction is also orthogonal to it.

Now we give some general results for a confounding arrangement. Suppose we wish to have a confounding arrangement in 2^p blocks of a 2^k factorial experiment. Then we have the following observations:

1. The size of each block is 2^{k-p}.
2. The number of elements in defining contrasts is $(2^p-1), i.e., (2^p-1)$ interactions have to be confounded.

   Proof: If p factors are to be confounded, then the number of mth order interaction with p factors is $\binom{p}{m}$, $(m=1,2,...,p)$. So the total number of factors to be confounded are $\sum_{m=1}^{p}\binom{p}{m}=2^{p-1}$.
3. If any two interactions are confounded, then their generalized interactions are also confounded.
4. The number of independent contrasts out of (2^p-1) defining contrasts is p and rest are obtained as generalized interactions.
5. Number of effects getting confounded automatically is (2^p-p-1).

To illustrate this, consider a 2^5 factorial $(k=5)$ with 5 factors, viz., A, B, C, D and E. The factors are to be confounded in 2^3 blocks ($p=3$). So the size of each block is $2^{5-3}=4$. The number of defining contrasts is $2^3-1=7$. The number of independent contrasts which can be chosen arbitrarily is 3(i.e., p) out of 7 defining contrasts. Suppose we choose $p=3$ following independent contrasts as

(i) ACE

(ii) CDE

(iii) $ABDE$

and then the remaining 4 out of 7 defining contrasts are obtained as

(iv) $(ACE)\times(CDE)=AC^2DE^2=AD$

(v) $(ACE)\times(ABDE)=A^2BCDE^2=BCD$

(vi) $(CDE)\times(ABDE)=ABCD^2E^2=ABC$

(vii) $(ACE)\times(CDE)\times(ABDE)=A^2BC^2D^2E^3=BE.$

Alternatively, if we choose another set of $p = 3$ independent contrast as

(i) $ABCD$,

(ii) $ACDE$,

(iii) $ABCDE$,

then the defining contrasts are obtained as

(iv) $(ABCD) \times (ACDE) = A^2BC^2D^2E = BE$

(v) $(ABCD) \times (ABCDE) = A^2B^2C^2D^2E = E$

(vi) $(ACDE) \times (ABCDE) = A^2BC^2D^2E^2 = B$

(vii) $(ABCD) \times (ACDE) \times (ABCDE) = A^3B^2C^3D^3E^2 = ACD.$

In this case, the main effects B and E also get confounded.

As a rule, try to confound, as far as possible, higher order interactions only because they are difficult to interpret.

After selecting p independent defining contrasts, divide the 2^k treatment combinations into 2^p groups of 2^{k-p} combinations each, and each group going into one block.

Principal (key) block:

Group containing the combination (1) is called the principal block or key block. It contains all the treatment combinations which are orthogonal to the chosen independent defining contrasts.

If there are p independent defining contrasts, then any treatment combination in principal block is orthogonal to p independent defining contrasts. In order to obtain the principal block,

- write the treatment combinations in standard order.
- check each one of them for orthogonality.
- If two treatment combinations belongs to the principal block, their product also belongs to the principal block.
- When few treatment combinations of the principal block have been determined, other treatment combinations can be obtained by multiplication rule.

Now we illustrate these steps in the following example.

Example:

Consider the set up of a 2^5 factorial experiment in which we want to divide the total treatment effects into 2^3 groups by confounding three effects AD, BE and ABC. The generalized interactions in this case are $ADBE$, BCD, ACE and CDE.

In order to find the principal block, first write the treatment combinations in standard order as follows:

(1)	a	b	ab	c	ac	bc	abc
d	ad	bd	abd	cd	acd	bcd	$abcd$
e	ae	be	abe	ce	ace	bce	$abce$
de	ade	bde	$abde$	cde	$acde$	$bcde$	$abcde$.

Place a treatment combination in the principal block if it has an even number of letters in common with the confounded effects AD, BE and ABC. The principal block has $(1), acd$, bce and $abde(= acd \times bce)$. Obtain other blocks of confounding arrangement from principal block by multiplying the treatment combinations of the principal block by a treatment combination not occurring in it or in any other block already obtained. In other words, choose treatment combinations not occurring in it and multiply with them in the principal block. Choose only distinct blocks. In this case, obtain other blocks by multiplying a, b, ab, c, ac, bc, abc like as in the following .

Arrangement of the treatments in blocks when AD, BE and ABC are confounded

Principal Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8
(1)	a	b	ab	c	ac	bc	abc
acd	cd	$abcd$	bcd	ad	d	abd	bd
bce	$abce$	ce	ace	be	abe	e	ae
$abde$	bde	ade	de	$abcde$	$bcde$	$acde$	cde

For example, block 2 is obtained by multiplying a with each factor combination in principal block as $(1) \times a = a$, $acd \times a = a^2cd = cd, bce \times a = abce$, $abde \times a = a^2bde = bde$; block 3 is obtained by multiplying b with (1), acd, bce and $abde$ and similarly other blocks are obtained. If any other treatment combination is chosen to be multiplied with the treatments in principal block, then we get

a block which will be one among the blocks 1 to 8. For example, if ae is multiplied with the treatments in principal block, then the blocks obtained consists of $(1)\times ae = ae$, $acd \times ae = cde$, $bce \times ae = abc$ and $abde \times ae = bd$ which is same as the block 8.

Alternatively, if ACD, $ABCD$ and $ABCDE$ are to be confounded, then independent defining contrasts are ACD, $ABCD$, $ABCDE$ and the principal block has (1), ac, ad and $cd(=ac\times ad)$.

Analysis of variance in case of confounded effects

When an effect is confounded, it means that it is not estimable. The following steps are followed to conduct the analysis of variance in case of factorial experiments with confounded effects:

- Obtain the sum of squares due to main and interaction effects in the usual way as if no effect is confounded.
- Drop the sum of squares corresponding to confounded effects and retain only the sum of squares due to unconfounded effects.
- Find the total sum of squares.
- Obtain the sum of squares due to error and associated degrees of freedom by subtraction.
- Conduct the test of hypothesis in the usual way.

Chapter 6

Balanced Incomplete Block Design (BIBD)

The designs like CRD and RBD are the complete block designs. We now discuss the balanced incomplete block design (BIBD) and the partially balanced incomplete block design (PBIBD) which are the incomplete block designs.

A balanced incomplete block design (BIBD) is an incomplete block design in which

- b blocks have the same number k of plots each and
- every treatment is replicated r times in the design.
- Each treatment occurs at most once in a block, i.e., $n_{ij} = 0$ or 1 where n_{ij} is the number of times the j^{th} treatment occurs in i^{th} block, $i = 1,2,...,b;\ j = 1,2,...,v.$
- Every pair of treatments occurs together is λ of the b blocks.

Five parameters denote such design as $D(b,k,v,r;\lambda)$.

The parameters b,k,v,r and λ are not chosen arbitrarily.

They satisfy the following relations:

(I) $bk = vr$

(II) $\lambda(v-1) = r(k-1)$

(III) $b \geq v$ (and hence $r > k$).

Hence $\sum_j n_{ij} = k$ for all i

$\sum_j n_{ij} = r$ for all j

and $n_{1j}n_{ij'} + n_{2j}n_{ij'} + ... + n_{b_j}n_{b_{j'}} = \lambda$ for all $j \neq j' = 1,2,...,v.$ Obviously $\frac{n_{ij}}{r}$ cannot be a constant for all j. So the design is not orthogonal.

Example of BIBD

In the design $D(b,k;v,r;\lambda)$: consider $b=10$ $(say, B_1,...,B_{10})$, $v=6$ $(say, T_1,...,T_6)$, $k=3, r=5, \lambda=2$

Blocks	Treatments		
B_1	T_1	T_2	T_5
B_2	T_1	T_2	T_6
B_3	T_1	T_3	T_4
B_4	T_1	T_3	T_6
B_5	T_1	T_4	T_5
B_6	T_2	T_3	T_4
B_7	T_2	T_3	T_5
B_8	T_2	T_4	T_6
B_9	T_3	T_5	T_6
B_{10}	T_4	T_5	T_6

Now we see how the conditions of BIBD are satisfied.

(i) $bk = 10\times 3 = 30$ and $vr = 6\times 5 = 30$

$\Rightarrow bk = vr$

(ii) $\lambda(v-1) = 2\times 5 = 10$ and $r(k-1) = 5\times 2 = 10$

$\Rightarrow \lambda(v-1) = r(k-1)$

(iii) $b = 10 \geq 6$

Even if the parameters satisfy the relations, it is not always possible to arrange the treatments in blocks to get the corresponding design.

The necessary and sufficient conditions to be satisfied by the parameters for the existence of a BIBD are not known.

The conditions (I)-(III) are some necessary condition only. The construction of such design depends on the actual arrangement of the treatments into blocks and this problem is handled in combinatorial mathematics. Tables are available, giving all the designs involving at most 20 replication and their method of construction.

Theorem:

$(I) bk = vr$

$(II)\ \lambda(v-1) = r(k-1)$

$(III)\ b \geq v.$

Proof: (I)

Let $N = (n_{ij}) : b \times v$ the incidence matrix

Observing that the quantities $E_{1b} N E_{v1}$ and $E_{1v} N' E_{b1}$ are the scalars and the transpose of each other, we find their values.

Consider

$$E_{1b} N E_{v1} = (1,1,...,1) \begin{pmatrix} n_{11} & n_{21} \cdots n_{b1} \\ n_{12} & n_{22} \cdots n_{b2} \\ \vdots & \vdots \ddots \vdots \\ n_{1v} & n_{2v} \cdots n_{bv} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$= (1,1,...,1) \begin{pmatrix} \sum_j n_{1j} \\ \sum_j n_{2j} \\ \vdots \\ \sum_j n_{bj} \end{pmatrix}$$

$$= (1,1,...,1)_{1\times b} \begin{pmatrix} k \\ k \\ \vdots \\ k \end{pmatrix} = bk.$$

Similarly,

$$E_{1v} N' E_{b1} = (1,...,1) \begin{pmatrix} n_{11} & n_{21} \cdots n_{b1} \\ n_{12} & n_{22} \cdots n_{b2} \\ \vdots & \vdots \ddots \vdots \\ n_{1v} & n_{2v} \cdots n_{bv} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ 1 \end{pmatrix}$$

$$= (1,1,...,1) \begin{pmatrix} \sum_i n_i \\ \vdots \\ \sum_i n_{iv} \end{pmatrix} = (1,1,...,1)_{1\times v} \begin{pmatrix} r \\ \vdots \\ r \end{pmatrix} = vr$$

But

$E_{1b} N E_{v1} = E_{1v} N' E_{b1}$ as both are scalars.

Thus $bk = vr$.

Proof: (II)

Consider

$$N'N = \begin{pmatrix} n_{11} & n_{21} \cdots n_{b1} \\ n_{12} & n_{22} \cdots n_{b2} \\ \vdots & \vdots \ddots \vdots \\ n_{1v} & n_{2v} \cdots n_{bv} \end{pmatrix} \begin{pmatrix} n_{11} & n_{12} \cdots n_{1v} \\ n_{21} & n_{22} \cdots n_{2v} \\ \vdots & \vdots \ddots \vdots \\ n_{b1} & n_{b2} \cdots n_{bv} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_i n_{i1}^2 & \sum_i n_{i1}n_{i2} \cdots \sum_i n_{i1}n_{iv} \\ \sum_i n_{i1}n_{i2} & \sum_i n_{i2}^2 \cdots \sum_i n_{i2}n_{iv} \\ \vdots & \vdots \ddots \vdots \\ \sum_i n_{iv}n_{i1} & \sum_i n_{iv}n_{i2} \cdots \sum_i n_{iv}^2 \end{pmatrix}$$

$$= \begin{pmatrix} r & \lambda & \cdots \lambda \\ \lambda & r & \cdots \lambda \\ \vdots & \vdots & \vdots \\ r & \lambda & \cdots \lambda \end{pmatrix}. \qquad (1)$$

Since $n_{ij}^2 = 1$ or 0 as $n_{ij} = 1$ or 0,

so $\sum_i n_{ij}^2$ = Number of times τ_j occurs in the design

$= r$ for all $j = 1, 2, ..., v$ of times occurs in the design

and $\sum_i n_{ij}n_{ij'}$ = Number of blocks in which τ_j and $\tau_{j'}$ occurs together

$= \lambda$ for all $j \neq j'$.

$$N'NE_{v1} = \begin{pmatrix} r & \lambda \cdots \lambda \\ \lambda & r \cdots \lambda \\ \vdots & \vdots \ddots \vdots \\ \lambda & \lambda \cdots r \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} r + \lambda(v-1) \\ r + \lambda(v-1) \\ \vdots \\ r + \lambda(v-1) \end{pmatrix} = [r + \lambda(v-1)]E_{v1}. \qquad (2)$$

Also

$$N'NE_{v1} = N'\left[\begin{pmatrix} n_{11} & n_{12}\cdots n_{1v} \\ n_{21} & n_{22}\cdots n_{2v} \\ \vdots & \vdots \ddots \vdots \\ n_{b1} & n_{b2}\cdots n_{bv} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}\right]$$

$$= N'\begin{pmatrix} \sum_j n_{1j} \\ \sum_j n_{2j} \\ \vdots \\ \sum_j n_{bj} \end{pmatrix}$$

$$= \begin{pmatrix} n_{11} & n_{21}\cdots n_{b1} \\ n_{12} & n_{22}\cdots n_{b2} \\ \vdots & \\ n_{iv} & n_{2v}\cdots n_{bv} \end{pmatrix}\begin{pmatrix} k \\ k \\ \vdots \\ k \end{pmatrix}_{b\times 1}$$

$$= k\begin{pmatrix} \sum_i n_{i1} \\ \sum_i n_{i2} \\ \vdots \\ \sum_i n_{iv} \end{pmatrix}$$

$$= k\begin{pmatrix} r \\ r \\ \vdots \\ r \end{pmatrix}$$

$$= krE_{v1} \qquad (3)$$

From (2) and (3)

$$[r+\lambda(v-1)]E_{v1} = krE_{v1}$$

or $\quad r+\lambda(v-1) = kr$

or $\quad \lambda(v-1) = r(k-1)$

Proof: (III)

From (I), the determinant of $N'N$ is

$$\det|N'N| = [r+\lambda(v-1)](r-\lambda)^{v-1}$$

$$= \left[r+r(k-1)\right](r-\lambda)^{v-1}$$

$$= rk(r-\lambda)^{v-1}$$

$$\neq 0$$

because since if $r = \lambda \Rightarrow$ from (II) that $k = v$. This contradicts the incompleteness of the design.

Thus $N'N$ is a $v \times v$ nonsingular matrix.

Thus $rank(N'N) = v.$

We know from matrix theory result

$rank(N) = rank(N'N)$

so $rank(N) = v$

But $rank(N) \le b$, there being b rows in N.

Thus $v \le b.$

Interpretation of conditions of BIBD

Interpretation of (I) *bk = vr*

This condition is related to the total number of plots in an experiment. In our settings, there are k plots in each block and there are b blocks. So the total number of plots are bk.

Further, there are v treatments and each treatment is replicated r times such that each treatment occurs atmost in one block. So total number of plots containing all the treatments is vr. Since both the statements counts the total number of plots, hence $bk = vr.$

Interpretation of (II)

Each block has k plots. Thus the total pairs of plots in a block = $\binom{k}{2} = \frac{k(k-1)}{2}.$

There are b blocks. Thus the total pairs of plots such that each pair consists of plots within a block = $b\frac{k(k-1)}{2}.$

There are v treatments, thus the total number of pairs of treatment = $\binom{v}{2} = \frac{v(v-1)}{2}.$

Each pair of treatment is replicated λ times, i.e., each pair of treatment occurs in λ blocks.

Thus the total number of pairs of plots within blocks must be $= \lambda\frac{v(v-1)}{2}.$

Hence $b\frac{k(k-1)}{2} = \lambda\frac{v(v-1)}{2}$

Using $bk = vr$ in this relation, we get $r(k-1) = \lambda(v-1).$

Proof of (III) was given by Fisher but quite long, so not needed here.

Balancing in designs:

There are two types of balancing – Variance balanced and efficiency balanced. We discuss the variance balancing now and the efficiency balancing later.

Balanced Design (Variance Balanced):

A connected design is said to be balanced (variance balanced) if all the elementary contrasts of the treatment effects can be estimated with the same precision. This definition does not hold for the disconnected design, as all the elementary contrasts are not estimable in this design.

Proper Design:

An incomplete block design with $k_1 = k_2 = ... = k_b = k$ is called a proper design.

Symmetric BIBD:

A BIBD is called symmetrical if the number of blocks = number of treatments, i.e., $b = v$.

Since $b = v$, so from $bk = vr$

$\Rightarrow k = r.$

Thus the number of pairs of treatments common between any two blocks = λ .

The determinant of $N'N$ is

$$\begin{aligned}|N'N| &= [r + \lambda(v-1)](r-\lambda)^{v-1} \\ &= \left[r + r(k-1)\right](r-\lambda)^{v-1} \\ &= rk(r-\lambda)^{v-1}.\end{aligned}$$

When BIBD is symmetric, $b = v$ and then using $bk = vr$, we have $k = r$. Thus

$|N'N| = |N|^2 = r^2(r-\lambda)^{v-1},$

so

$|N| = \pm\, r(r-\lambda)^{\frac{v-1}{2}}.$

Since $|N|$ is an integer, hence when v is an even number, $(r-\lambda)$ must be a perfect square. So

$$\begin{aligned}|N'N| &= (r-\lambda)I + \lambda E_{v1}E_{v1}', \\ (N'N)^{-1} &= N^{-1}N'^{-1} \\ &= \frac{1}{r-\lambda}\left[I - \frac{\lambda}{r^2}E_{v1}E_{v1}'\right], \\ N'^{-1} &= \frac{1}{r-\lambda}\left[I - \frac{\lambda}{r^2}E_{v1}E_{v1}'\right].\end{aligned}$$

Post-multiplying both sides by N', we get

$$NN' = (r-\lambda)I + \lambda E_{v1}E_{v1}' = N'N.$$

Hence in the case of a symmetric BIBD, any two blocks have λ treatment in common.

Since BIBD is an incomplete block design. So every pair of treatment can occur at most once is a block, we must have $v \geq k$.

If $v = k$, then it means that each treatment occurs once in every block which occurs in case of RBD. So in BIBD, always assume $v > k$.

Similarly $\lambda < r$.

[If $\lambda = r$ then $\lambda(v-1) = r(k-1) \Rightarrow v = k \Rightarrow$ which means that the design is RBD]

Resolvable design:

A block design of

- b blocks in which
- each of v treatments is replicated r times

is said to be resolvable if b blocks can be divided into r sets of b/r blocks each, such that every treatment appears in each set precisely once. Obviously, in a resolvable design, b is a multiple of r.

Theorem: If in a BIBD $D(v,b,r,k,\lambda)$, b is divisible by r, then

$$b \geq v + r - 1.$$

Proof: Let $b = nr$ (where $n > 1$ is a positive integer).

For a BIBD, $\lambda(v-1) = r(k-1)$

or $r = \dfrac{\lambda(v-1)}{(k-1)}$ $\quad \left[\begin{array}{ll} \text{because} & vr = bk \\ \text{or} & vr = nrk \\ \text{or} & v = nk \end{array}\right]$

$$= \frac{\lambda(nk-1)}{(k-1)}$$

$$= \lambda\left(\frac{n-1}{k-1}\right) + \lambda n.$$

Since $n > 1$ and $k > 1$, so $\lambda n > 1$ is an integer. Since r has to be an integer.

$\Rightarrow \lambda\dfrac{(n-1)}{k-1}$ is also a positive integer.

Now, if possible, let

$$b < v + r - 1$$
$$\Rightarrow nr < v + r - 1$$
$$\text{or } r(n-1) < v - 1$$
$$\text{or } r(n-1) < \frac{r(k-1)}{\lambda} \qquad (\text{because } v - 1 = \frac{r(k-1)}{\lambda})$$
$$\Rightarrow \frac{\lambda(n-1)}{k-1} < 1 \text{ which is a contradiction as integer can not be less than one}$$
$$\Rightarrow b < v + r - 1 \text{ is impossible. Thus the opposite is true.}$$
$$\Rightarrow b \geq v + r - 1 \text{ holds correct.}$$

Intrablock analysis of BIBD:

Consider the model

$$y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij};\ i = 1,2,...,b;\ j = 1,2,...,v,$$

where

μ is the general mean effect;

β_i is the fixed additive i^{th} block effect;

τ_j is the fixed additive j^{th} treatment effect and

ε_{ij} is the i.i.d. random error with $\varepsilon_{ij} \sim N(0, \sigma^2)$.

We don't need to develop the analysis of BIBD from starting. Since BIBD is also an incomplete block design and the analysis of incomplete block design has already been presented in the earlier module, so we implement those derived expressions directly under the setup and conditions of BIBD. Using the same notations, we represent the blocks totals by $B_i = \sum_{j=1}^{v} y_{ij}$, treatment totals by $V_j = \sum_{i=1}^{b} y_{ij}$, adjusted treatment totals by Q_j and grand total by $G = \sum_{i=1}^{b}\sum_{j=}^{v} y_{ij}$ The normal equations are obtained by differentiating the error sum of squares. Then the block effects are eliminated from the normal equations and the normal equations are solved for the treatment effects. The resulting intrablock equations of treatment effects in matrix notations are expressible as

$$Q = C\hat{\tau}.$$

Now we obtain the forms of C and Q in the case of BIBD. The diagonal elements of C are given by

$$c_{jj} = r - \frac{\sum_{i=1}^{b} n_{ij}^2}{k} \quad (j = 1,2,...,v)$$
$$= r - \frac{r}{k}.$$

The off-diagonal elements of C are given by

$$c_{jj'} = -\frac{1}{k}\sum_{i=1}^{b} n_{ij}n_{ij'} \quad (j \neq j'; j, j' = 1,2,...,v)$$
$$= -\frac{\lambda}{k}.$$

The adjusted treatment totals are obtained as

$$Q_j = V_j - \frac{1}{k}\sum_{i=1}^{b} n_{ij}B_i \quad (j \neq 1,2,...,v)$$
$$= V_j - \frac{1}{k}\sum_{i(j)} B_i$$

where $\sum_{i(j)}$ denotes the sum over those blocks containing j^{th} treatment. Denote

$T_j = \sum_{i(j)} B_i$, then

$$Q_j = V_j - \frac{T_j}{k}.$$

The C matrix is simplified as follows:

$$C = rI - \frac{N'N}{k}$$
$$= rI - \frac{1}{k}\left[(r-\lambda)I + \lambda E_{v1}E_{v1}'\right]$$
$$= r\left(\frac{k-1}{k}\right)I + \frac{\lambda}{k}(I - E_{v1}E_{vi}')$$
$$= \lambda\left(\frac{v-1}{k}\right)I + \frac{\lambda}{k}(I - E_{v1}E_{vi}')$$
$$= \frac{\lambda V}{k}\left(I - \frac{E_{v1}E_{v1}'}{v}\right).$$

Since C is not as a full rank matrix, so its unique inverse does not exist. The generalized inverse of C is denoted as C^- which is obtained as

$$C^- = \left(C + \frac{E_{v1}E_{v1}'}{v}\right)^{-1}.$$

Since

$$C = \frac{\lambda v}{k}\left(I_v - \frac{E_{v1}E_{v1}'}{v}\right)$$

or $\frac{kC}{\lambda v} = I_v - \frac{E_{v1}E_{v1}'}{v}$,

the generalized inverse of $\frac{k}{\lambda v}C$ is

$$\left(\frac{k}{\lambda v}\right)^{-1} C^{-} = \left[C + \frac{E_{v1}E_{v1}^{'}}{v},\right]^{-1}$$

$$= \left[I_v - \frac{E_{v1}E_{v1}^{'}}{v} + \frac{E_{v1}E_{v1}^{'}}{v}\right]^{-1}$$

$$= I_v.$$

Thus $C^{-} = \frac{\lambda v}{k} I_v$.

Thus an estimate of τ is obtained from $Q = C\tau$ as

$$\hat{\tau} = C^{-}Q$$

$$= \frac{\lambda v}{k} Q.$$

The null hypothesis of our interest is $H_0 : \tau_1 = \tau_2 = ... = \tau_v$ against the alternative hypothesis H_1 : at least one pair of τ_j's is different. Now we obtain the various sum of squares involved in the development of analysis of variance as follows.

The adjusted treatment sum of squares is

$$SS_{Treat(adj)} = \hat{\tau}'Q$$

$$= \frac{k}{\lambda v} Q'Q$$

$$= \frac{k}{\lambda v} \sum_{j=1}^{v} Q_j^2,$$

The unadjusted block sum of squares is

$$SS_{Block(unadj)} = \sum_{i=1}^{b} \frac{B_i^2}{k} - \frac{G^2}{bk} \ .$$

The total sum of squares is

$$SS_{Total} = \sum_{i=1}^{b} \sum_{j=1}^{v} y_{ij}^2 - \frac{G^2}{bk}$$

The residual sum of squares is obtained by

$$SS_{Error(t)} = SS_{Total} - SS_{Block(unadj)} - SS_{Treat(adj)}.$$

A test for $H_0 : \tau_1 = \tau_2 = ... = \tau_v$ is then based on the statistic

$$F_{Tr} = \frac{SS_{Treat(adj)}/(v-1)}{SS_{Error(t)}/(bk-b-v+1)}$$

$$= \frac{k}{\lambda v} . \frac{bk-b-v+1}{v-1} . \frac{\sum_{j=1}^{v} Q_j^2}{SS_{Error(t)}}$$

If $F_{Tr} > F_{1-\alpha, v-1, bk-b-v+1;}$ then $H_{0(t)}$ is rejected.

This completes the analysis of variance test and is termed as intrablock analysis of variance. This analysis can be compiled into the intrablock analysis of variance table for testing the significance of the treatment effect given as follows.

Intrablock analysis of variance table of BIBD for $H_0 : \tau_1 = \tau_2 = ... = \tau_v$

Source	Sum of squares	Degrees of freedom	Mean squares	F
Between treatment (adjusted)	$SS_{Treat(adj)}$	$v-1$	$MS_{treat} = \frac{SS_{Treat(adj)}}{v-1}$	$\frac{MS_{Treat}}{MS_E}$
Between blocks (unadjusted)	$SS_{Block(unadj)}$	b - 1		
Intrablock error	$SS_{Error(t)}$ (by substraction)	$bk - b - v + 1$	$MS_E = \frac{SS_{Error(t)}}{bk-b-v+1}$	
Total	$SS_{Total} = \sum_i \sum_j y_{ij}^2 - \frac{G^2}{bk}$	$bk-1$		

In case, the null hyperthesis is rejected, then we go for a pairwise comparison of the treatments. For that, we need an expression for the variance of the difference of two treatment effects.

The variance of an elementary contrast $(\tau_j - \tau_{j'}, j \neq j')$ under the intrablock analysis is

$$\begin{aligned}
V^* &= Var(\hat{\tau}_j - \hat{\tau}_{j'}) \\
&= Var\left(\frac{k}{\lambda v}(Q_j - Q_{j'})\right) \\
&= \frac{k^2}{\lambda^2 v^2}[Var(Q_j) + Var(Q_j) - 2Cov(Q_j Q_{j'})] \\
&= \frac{k^2}{\lambda^2 v^2}(c_{jj} + c_{j'j'} - 2c_{jj'})\sigma^2 \\
&= \frac{k^2}{\lambda^2 v^2}\left[2r\left(1 - \frac{1}{k}\right) + \frac{2\lambda}{k}\right]\sigma^2 \\
&= \frac{2k}{\lambda v}\sigma^2.
\end{aligned}$$

This expression depends on σ^2 which is unknown. So it is unfit for use in the real data applications. One solution is to estimate σ^2 from the given data and use it in the place of σ^2.

An unbiased estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{SS_{Error(t)}}{bk - b - v + 1}.$$

Thus an unbiased estimator of V^* can be obtained by substituting $\hat{\sigma}^2$ in it as

$$\hat{V}_* = \frac{2k}{\lambda v}.\frac{SS_{Error(t)}}{bk - b - v + 1}.$$

If H_0 is rejected, then we make pairwise comparison and use the multiple comparison test. To test $H_0 : \tau_j = \tau_{j'} (j \neq j')$, a suitable statistic is

$$t = \frac{k(bk - b - v + 1)}{\lambda v}.\frac{Q_j - Q_{j'}}{\sqrt{SS_{Error(t)}}}$$

which follows a *t*-distribution with $(bk - b - v + 1)$ degrees of freedom under H_0.

A question arises that how a BIBD compares to an RBD. Note that BIBD is an incomplete block design whereas RBD is a complete block design. This point should be kept in mind while making such restrictive comparison.

We now compare the efficiency of BIBD with a randomized block (complete) design with r replicates. The variance of an elementary contrast under a randomized block design (RBD) is

$$V_R^* = Var(\hat{\tau}_j^2 - \hat{\tau}_{j'})_{RBD} = \frac{2\sigma_*^2}{r}$$

where $Var(y_{ij}) = \sigma_*^2$ under RBD.

Thus the relative efficiency of BIBD relative to RBD is

$$\frac{Var(\hat{\tau}_j - \hat{\tau}_{j'})_{RBD}}{Var(\hat{\tau}_j - \hat{\tau}_{j'})_{BIBD}} = \frac{\left(\frac{2\sigma_*^2}{r}\right)}{\left(\frac{2k\sigma^2}{\lambda v}\right)}$$

$$= \frac{\lambda v}{rk}\left(\frac{\sigma_*^2}{\sigma^2}\right).$$

The factor $\frac{\lambda v}{rk} = E$ (say) is termed as the **efficiency factor** of BIBD and

$$E = \frac{\lambda v}{rk} = \frac{v}{k}\left(\frac{k-1}{v-1}\right)$$

$$= \left(1 - \frac{1}{k}\right)\left(1 - \frac{1}{v}\right)^{-1}$$

$$< 1 \text{ (since } v > k).$$

The actual efficiency of BIBD over RBD not only depends on the efficiency factor but also on the ratio of variances σ_*^2 / σ^2. So BIBD can be more efficient than RBD as σ_*^2 can be more than σ^2 because $k < v$.

Efficiency balanced design:

A block design is said to be efficiency balanced if every contrast of the treatment effects is estimated through the design with the same efficiency factor.

If a block design satisfies any two of the following properties:

(i) efficiency balanced,

(ii) variance balanced and

(iii) an equal number of replications,

then the third property also holds true.

Missing observations in BIBD:

The intrablock estimate of missing $(i, j)^{th}$ observation y_{ij} is

$$y_{ij} = \frac{vr(k-1)B_i - k(v-1)Q_j - (v-1)Q_j'}{k(k-1)(bk - b - v + 1)}$$

Q_j' : the sum of Q value for all other treatment (but not the j^{th} one) which are present in the i^{th} block.

All other procedures remain the same.

Interblock analysis and recovery of interblock information in BIBD

In the intrablock analysis of variance of an incomplete block design or BIBD, the treatment effects were estimated after eliminating the block effects from the normal equations. In a way, the block effects were assumed to be not marked enough and so they were eliminated. It is possible in many situations that the block effects are influential and marked. In such situations, the block totals may carry information about the treatment combinations also. This information can be used in estimating the treatment effects which may provide more efficient results. This is accomplished by an interblock analysis of BIBD and used further through the recovery of interblock information. So we first conduct the interblock analysis of BIBD. We do not derive the expressions a fresh but we use the assumptions and results from the interblock analysis of an incomplete block design. We additionally assume that the block effects are random with variance σ_β^2.

After estimating the treatment effects under interblock analysis, we use the results for the pooled estimation and recovery of interblock information in a BIBD.

In case of BIBD,

$$N'N = \begin{pmatrix} \sum_i n_{i1}^2 & \sum_i n_{i1}n_{i2} & \cdots & \sum_i n_{i1}n_{iv} \\ \sum_i n_{i1}n_{i2} & \sum_i n_{i2}^2 & \cdots & \sum_i n_{i2}n_{iv} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_i n_{iv}n_{i1} & \sum_i n_{iv}n_{i2} & \cdots & \sum_i n_{iv}^2 \end{pmatrix}$$

$$= \begin{pmatrix} r & \lambda & \cdots & \lambda \\ \lambda & r & \cdots & \lambda \\ \vdots & \vdots & \ddots & \vdots \\ \lambda & \lambda & \cdots & r \end{pmatrix}$$

$$= (r-\lambda)I_v + \lambda E_{v1}E_{v1}'$$

$$(N'N)^{-1} = \frac{1}{r-\lambda}\left[I_v - \frac{\lambda E_{v1}E_{v1}'}{rk}\right]$$

The interblock estimate of τ can be obtained by substituting the expression on $(N'N)^{-1}$ in the earlier obtained interblock estimate.

$$\tilde{\tau} = (N'N)^{-1}N'B - \frac{GE_{v1}}{bk}.$$

Our next objective is to use the intrablock and interblock estimates of treatment effects together to find an improved estimate of treatment effects.

In order to use the interblock and intrablock estimates of τ together through pooled estimate, we consider the interblock and intrablock estimates of the treatment contrast.

The intrablock estimate of treatment contrast $l'\tau$ is

$$\begin{aligned} l'\hat{\tau} &= l'C^{-}Q \\ &= \frac{k}{\lambda v} l'Q \\ &= \frac{k}{\lambda v} \sum_j l_j Q_j \\ &= \sum_j l_j \hat{\tau}_j, \text{ say.} \end{aligned}$$

The interblock estimate of treatment contrast $l'\tau$ is

$$\begin{aligned} l'\tilde{\tau} &= \frac{l'N'B}{r-\lambda} \quad (\text{since } l'E_{v1} = 0) \\ &= \frac{1}{r-\lambda} \sum_{j=1}^{v} l_j \left(\sum_{i=1}^{b} n_{ij} B_i \right) \\ &= \frac{1}{r-\lambda} \sum_{j=1}^{v} l_j T_j \\ &= \sum_{j=1}^{v} l_j \tilde{\tau}_j. \end{aligned}$$

The variance of $l'\hat{\tau}$ is obtained as

$$\begin{aligned} Var(l'\hat{\tau}) &= \left(\frac{k}{\lambda v} \right) Var\left(\sum_j l_j Q_j \right) \\ &= \left(\frac{k}{\lambda v} \right)^2 \left[\sum_j l_j^2 Var(Q_j) + 2 \sum_j \sum_{j'(\neq j)} l_j l_{j'} Cov(Q_j, Q_{j'}) \right]. \end{aligned}$$

Since

$$Var(Q_j) = r\left(1 - \frac{1}{k}\right)\sigma^2,$$

$$Cov(Q_j, Q_{j'}) = -\frac{\lambda}{k}\sigma^2, \quad (j \neq j'),$$

so

$$Var(l'\hat{\tau}) = \left(\frac{k}{\lambda v}\right)^2 \left[r\left(1-\frac{1}{k}\right)\sigma^2 \sum_j l_j^2 - \frac{\lambda}{k}\left\{ \left(\sum_j l_j\right)^2 - \sum_j l_j^2 \right\}\sigma^2 \right]$$

$$= \left(\frac{k}{\lambda v}\right)^2 \left[\frac{r(k-1)}{k} \sum_j l_j^2 + \frac{\lambda}{k} \sum_j l_j^2 \right]\sigma^2 \text{ (since } \sum_j \ell_j = 0 \text{ being contrast)}$$

$$= \left(\frac{k}{\lambda v}\right)^2 \frac{1}{k}\left[\lambda(v-1)+\lambda\right]\sum_j l_j^2 \qquad \text{(using } r(k-1) = \lambda(v-1)\text{)}$$

$$= \left(\frac{k}{\lambda v}\right)\sigma^2 \sum_j l_j^2.$$

Similarly, the variance of $\ell'\hat{\tau}$ is obtained as

$$Var(l'\tilde{\tau}) = \left(\frac{1}{r-\lambda}\right)^2 \left[\sum_j l_j^2 Var(T_j) + 2\sum_j \sum_{j'(\neq j)} l_j l_{j'} Cov(T_j, T_{j'}) \right]$$

$$= \left(\frac{1}{r-\lambda}\right)^2 \left[r\sigma_f^2 \sum_j l_j^2 + \lambda\sigma_f^2 \left\{ \left(\sum_j l_j\right)^2 - \sum_j l_j^2 \right\} \right]$$

$$= \frac{\sigma_f^2}{r-\lambda} \sum_j l_j^2.$$

The information on $\ell'\hat{\tau}$ and $\ell'\tilde{\tau}$ can be used together to obtain a more efficient estimator of $\ell'\tau$ by considering the weighted arithmetic mean of $\ell'\hat{\tau}$ and $\ell'\tilde{\tau}$. This will be the minimum variance unbiased and estimator of $\ell'\tau$ when the weights of the corresponding estimates are chosen such that they are inversely proportional to the respective variances of the estimators. Thus the weights to be assigned to intrablock and interblock estimates are reciprocal to their variances as $\lambda v/(k\sigma^2)$ and $(r-\lambda)/\sigma_f^2$, respectively. Then the pooled mean of these two estimators is

$$L^* = \frac{\frac{\lambda v}{k\sigma^2} l'\hat{\tau} + \frac{r-\lambda}{\sigma_f^2} l'\hat{\tau}}{\frac{\lambda v}{k\sigma^2} + \frac{r-\lambda}{\sigma_f^2}} = \frac{\frac{\lambda v}{k\sigma^2}\sum_j l_j\hat{\tau}_j + \frac{r-\lambda}{\sigma_f^2}\sum_j l_j\tilde{\tau}_j}{\frac{\lambda v}{k\sigma^2} + \frac{r-\lambda}{\sigma_f^2}}$$

$$= \frac{\frac{\lambda v\omega_1}{k}\sum_j l_j\hat{\tau}_j + (r-\lambda)\omega_2 \sum_j l_j\tilde{\tau}_j}{\frac{\lambda v}{k}\omega_1 + (r-\lambda)\omega_2}$$

$$= \frac{\lambda v\omega_1 \sum_j l_j\hat{\tau}_j + k(r-\lambda)\omega_2 \sum_j l_j\tilde{\tau}_j}{\lambda v\omega_1 + k(r-\lambda)\omega_2}$$

$$= \sum_j l_j \left[\frac{\lambda v\omega_1\hat{\tau}_j + k(r-\lambda)\omega_2\tilde{\tau}_j}{\lambda v\omega_1 + k(r-\lambda)\omega_2} \right]$$

$$= \sum_j l_j \tau_j^*$$

where $\tau_j^* = \dfrac{\lambda v \omega_1 \hat{\tau}_j + k(r-\lambda)\omega_2 \hat{\tau}_j}{\lambda v \omega_1 + k(r-\lambda)\omega_2}$, $\omega_1 = \dfrac{1}{\sigma^2}, \omega_2 = \dfrac{1}{\sigma_f^2}$.

Now we simplify the expression of τ_j^* so that it becomes more compatible in further analysis.

Since $\hat{\tau}_j = (k/\lambda v)Q_j$ and $\tilde{\tau}_j = T_j/(r-\lambda)$, so the numerator of τ_j^* can be expressed as

$$\omega_1 \lambda v \hat{\tau}_j + \omega_2 k(r-\lambda)\tilde{\tau}_j = \omega_1 k Q_j + \omega_2 k T_j$$

Similarly, the denominator of τ_j^* can be expressed as

$$\begin{aligned}
&\omega_1 \lambda v + \omega_2 k(r-\lambda) \\
&= \omega_1 \left[\frac{vr(k-1)}{v-1}\right] + \omega_2 \left[k\left(r - \frac{r(k-1)}{v-1}\right)\right] \qquad (\text{using } \lambda(v-1) = r(k-1)) \\
&= \frac{1}{v-1}\left[\omega_1 vr(k-1) + \omega_2 kr(v-k)\right].
\end{aligned}$$

Let

$$W_j^* = (v-k)V_j - (v-1)T_j + (k-1)G$$

where $\sum_j W_j^* = 0$. Using these results we have

$$\begin{aligned}
\tau_j^* &= \frac{(v-1)\left[\omega_1 k Q_j + \omega_2 k T_j\right]}{\omega_1 rv(k-1) + \omega_2 kr(v-k)} \\
&= \frac{(v-1)\left[\omega_1 (kV_j - T_j) + \omega_2 k T_j\right]}{r\left[\omega_1 v(k-1) + \omega_2 k(v-k)\right]} \quad (\text{using } Q_j = V_j - \frac{T_j}{k}) \\
&= \frac{\omega_1 k(v-1)V_j + (k\omega_2 - \omega_1)(v-1)T_j}{r\left[\omega_1 v(k-1) + \omega_2 k(v-k)\right]} \\
&= \frac{\omega_1 k(v-1)V_j + (\omega_1 - k\omega_2)\left[W_j^* - (v-k)V_j - (k-1)G\right]}{r\left[\omega_1 v(k-1) + \omega_2 k(v-k)\right]} \\
&= \frac{\left[\omega_1 k(v-1) - (\omega_1 - k\omega_2)(v-k)\right]V_j + (\omega_1 - k\omega_2)\left[W_j^* - (k-1)G\right]}{r\left[\omega_1 v(k-1) + \omega_2 k(v-k)\right]} \\
&= \frac{1}{r}\left[V_j + \frac{\omega_1 - k\omega_2}{\omega_1 v(k-1) + \omega_2 k(v-k)}\left\{W_j^* - (k-1)G\right\}\right] \\
&= \frac{1}{r}\left[V_j + \xi\left\{W_j^* - (k-1)G\right\}\right]
\end{aligned}$$

where

$$\xi = \frac{\omega_1 - k\omega_2}{\omega_1 v(k-1) + \omega_2 k(v-k)}, \omega_1 = \frac{1}{\sigma^2}, \omega_2 = \frac{1}{\sigma_f^2}.$$

Thus the pooled estimate of the contrast $l'\tau$ is

$$l'\tau^* = \sum_j l_j \tau_j^*$$

$$= \frac{1}{r}\sum_j l_j (V_j + \xi W_j^*) \quad (\text{since } \sum_j l_j = 0 \text{ being contrast})$$

The variance of $l'\tau^*$ is

$$Var(l'\tau^*) = \frac{k}{\lambda v\omega_1 + k(r-\lambda)\omega_2}\sum_j l_j^2$$

$$= \frac{k(v-1)}{r\left[v(k-1\omega_1 + k(v-k)\omega_2\right]}\sum_j l_j^2 \qquad (\text{using } \lambda(v-1) = r(k-1)$$

$$= \sigma_E^2 \frac{\sum_j l_j^2}{r}$$

where

$$\sigma_E^2 = \frac{k(v-1)}{v(k-1)\omega_1 + k(v-k)\omega_2}$$

is called as the **effective variance**.

Note that the variance of any elementary contrast based on the pooled estimates of the treatment effects is

$$Var(\tau_i^* - \tau_j^*) = \frac{2}{r}\sigma_E^2.$$

The effective variance can be approximately estimated by

$$\hat{\sigma}_E^2 = MSE\left[1 + (v-k)\omega^*\right]$$

where MSE is the mean square due to error obtained from the intrablock analysis as

$$MSE = \frac{SS_{Error(t)}}{bk - b - v + 1}$$

and

$$\omega^* = \frac{\omega_1 - \omega_2}{v(k-1)\omega_1 + k(v-k)\omega_2}.$$

The quantity ω^* depends upon the unknown σ^2 and σ_β^2. To obtain an estimate of ω^*, we can obtain the unbiased estimates of σ^2 and σ_β^2 and then substitute them back in place of σ^2 and σ_β^2 in ω^*. To do this, we proceed as follows.

An estimate of ω_1 can be obtained by estimating σ^2 from the intrablock analysis of variance as

$$\hat{\omega}_1 = \frac{1}{\hat{\sigma}_2} = [MSE]^{-1}.$$

The estimate of ω_2 depends on $\hat{\sigma}^2$ and $\hat{\sigma}_\beta^2$. To obtain an unbiased estimator of σ_β^2, consider

$$SS_{Block(adj)} = SS_{Treat(adj)} + SS_{Block(unadj)} - SS_{Treat(unadj)}$$

for which

$$E(SS_{Block(adj)}) = (bk - v)\sigma_\beta^2 + (b-1)\sigma^2.$$

Thus an unbiased estimator of σ_β^2 is

$$\begin{aligned}
\hat{\sigma}_\beta^2 &= \frac{1}{bk-v}\left[SS_{Block(adj)} - (b-1)\hat{\sigma}^2\right] \\
&= \frac{1}{bk-v}\left[SS_{Block(adj)} - (b-1)MSE\right] \\
&= \frac{b-1}{bk-v}\left[MS_{Block(adj)} - MSE\right] \\
&= \frac{b-1}{v(r-1)}\left[MS_{Block(adj)} - MSE\right]
\end{aligned}$$

where

$$MS_{Block(adj)} = \frac{SS_{Block(adj)}}{b-1}.$$

Thus

$$\begin{aligned}
\hat{\omega}_2 &= \frac{1}{k\hat{\sigma}^2 + \hat{\sigma}_\beta^2} \\
&= \frac{1}{v(r-1)\left[k(b-1)SS_{Block(adj)} - (v-k)SS_{Error(t)}\right]}.
\end{aligned}$$

Recall that our main objective is to develop a test of hypothesis for $H_0 : \tau_1 = \tau_2 = ... = \tau_v$ and we now want to develop it using the information based on both interblock and intrablock analysis.

To test the hypothesis related to treatment effects based on the pooled estimate, we proceed as follows.

Consider the adjusted treatment totals based on the intrablock and the interblock estimates as

$$T_j^* = T_j + \omega * W_j^*; \; j = 1, 2, ..., v$$

and use it as usual treatment total as in earlier cases.

The sum of squares due to T_j^* is

$$S_{T^*}^2 = \sum_{j=1}^{v} T_j^{*2} - \frac{\left(\sum_{j=1}^{v} T_j^*\right)^2}{v}.$$

Note that in the usual analysis of variance technique, the test statistic for such hull hypothesis is developed by taking the ratio of the sum of squares due to treatment divided by its degrees of freedom and the sum of squares due to error divided by its degrees of freedom. Following the same idea, we define the statistics

$$F^* = \frac{S_{T^*}^2/[(v-1)r]}{MSE[1+(v-k)\hat{\omega}^*]}$$

where $\hat{\omega}^*$ is an estimator of ω^*. It may be noted that F^* depends on $\hat{\omega}^*$. The value of $\hat{\omega}^*$ itself depends on the estimated variances $\hat{\sigma}^2$ and $\hat{\sigma}_f^2$. So it cannot be ascertained that the statistic F^* necessary follow the F distribution. Since the construction of F^* is based on the earlier approaches where the statistic was found to follow the exact F -distribution, so based on this idea, the distribution of F^* can be considered to be approximately F distributed. Thus the approximate distribution of F^* is considered as F distribution with $(v-1)$ and $(bk-b-v+1)$ degrees of freedom. Also, $\hat{\omega}^*$ is an estimator of ω^* which is obtained by substituting the unbiased estimators of ω_1 and ω_2.

An approximate best pooled estimator of $\sum_{j=1} l_j \tau_j$ is

$$\sum_{j=1}^{v} l_j \frac{V_j + \hat{\xi} W_j}{r}$$

and its variance is approximately estimated by

$$\frac{k\sum_j l_j^2}{\lambda v\hat{\omega}_1 + (r-\lambda)k\hat{\omega}_2}.$$

In case of the resolvable BIBD, $\hat{\sigma}_\beta^2$ can be obtained by using the adjusted block with replications sum of squares from the intrablock analysis of variance. If sum of squares due to such block total is SS_{Block}^* and corresponding mean square is

$$MS_{Block}^* = \frac{SS_{Block}^*}{b-r}$$

then

$$E(MS^*_{Block}) = \sigma^2 + \frac{(v-k)(r-1)}{b-r}\sigma^2_\beta$$
$$= \sigma^2 + \frac{(r-1)k}{r}\sigma^2_\beta$$

and $k(b-r) = r(v-k)$ for a resolvable design. Thus

$$E\left[rMS^*_{Block} - MSE\right] = (r-1)(\sigma^2 + k\sigma^2_\beta)$$

and hence

$$\hat{\omega}_2 = \left[\frac{rMS^*_{block} - MSE}{r-1}\right]^{-1},$$
$$\hat{\omega}_1 = \left[MSE\right]^{-1}.$$

The analysis of variance table for the recovery of interblock information in BIBD is described in the following table:

Source	Sum of squares	Degrees of freedom	Mean square	F^*
Between treatment (unadjusted)	$S^2_{T^*}$	v - 1		$F^* = \frac{MS_{Blocks(adj)}}{MSE}$
Between blocks (adjusted)	$SS_{Block(adj)} = SS_{Treat(adj)} + SS_{Block(unadj)} - SS_{Treat(unadj)}$	b - 1	$MS_{Blocks(adj)} = \frac{SS_{Block(adj)}}{b-1}$	
Intrablock error	$SS_{Error(t)}$ (by substraction)	$bk - b - v + 1$	$MSE = \frac{SS_{Error(t)}}{bk-b-v+1}$	
Total	SS_{Total}	bk - 1		

The increase in the precision using interblock analysis as compared to intrablock analysis is

$$\frac{Var(\hat{\tau})}{Var(\tau^*)} - 1$$
$$= \frac{\lambda v \omega_1 + \omega_2 k(r-\lambda)}{\lambda v \omega_1} - 1$$
$$= \frac{\omega_2 (r-\lambda) k}{\lambda v \omega_1}.$$

Such an increase may be estimated by

$$\frac{\hat{\omega}_2 (r-\lambda) k}{\lambda v \hat{\omega}_1}.$$

Although $\omega_1 > \omega_2$ but this may not hold true for $\hat{\omega}_1$and $\hat{\omega}_2$. The estimates $\hat{\omega}_1$ and $\hat{\omega}_2$ may be negative also and in that case we take $\hat{\omega}_1 = \hat{\omega}_2$.

SPLIT PLOT AND STRIP PLOT DESIGNS

D.K. Sehgal
I.A.S.R.I., Library Avenue, New Delhi- 110 012

1. Split Plot Design

1.1 Introduction

In conducting experiments, sometimes some factors have to be applied in larger experimental units while some other factors can be applied in comparatively smaller experimental units. Further some experimental materials may be rare while the other experimental materials may be available in large quantity or when the levels of one (or more) treatment factors are easy to change, while the alteration of levels of other treatment factors are costly, or time-consuming. One more point may be that although two or more different factors are to be tested in the experiment, one factor may require to be tested with higher precision than the others. In all such situations, a design called the split plot design is adopted.

A split plot design is a design with at least one blocking factor where the experimental units within each block are assigned to the treatment factor levels as usual, and in addition, the blocks are assigned at random to the levels of a further treatment factor. The designs have a nested blocking structure. In a block design, the experimental units are nested within the blocks, and a separate random assignment of units to treatments is made within each block. In a split plot design, the experimental units are called split-plots (or sub-plots), and are nested within whole plots (or main plots).

In split plot design, plot size and precision of measurement of effects are not the same for both factors, the assignment of a particular factor to either the main plot or the sub-plot is extremely important. To make such a choice, the following guidelines are suggested:

Degree of Precision- For a greater degree of precision for factor B than for factor A, assign factor B to the sub-plot and factor A to the main plot e.g. a plant breeder who plans to evaluate ten promising rice varieties with three levels of fertilization, would probably wish to have greater precision for varietal comparison than for fertilizer response. Thus, he would designate variety as the sub-plot factor and fertilizer as the main plot factor. Or, an agronomist would assign variety to main plot and fertilizer to sub-plot if he wants greater precision for fertilizer response than variety effect.

Relative Size of the Main effects- If the main effect of one factor (A) is expected to be much larger and easier to detect than that of the other factor (B), factor A can be assigned to the main plot and factor B to the sub-plot. This increases the chance of detecting the difference among levels of factor B which has a smaller effect.

Management Practices- The common type of situation when the split plot design is automatically suggestive is the difficulties in the execution of other designs, i.e. practical execution of plans. The cultural practices required by a factor may dictate the use of large

plots. For practical expediency, such a factor may be assigned to the main plot e.g. in an experiment to evaluate water management and variety, it may be desirable to assign water management to the main plot to minimize water movement between adjacent plots, facilitate the simulation of the water level required, and reduce border effects. Or, if ploughing is one of the factors of interest, then one cannot have different depths of ploughing in different plots scattered randomly apart.

1.2 Randomization and Layout

There are two separate randomization processes in a split plot design – one for the main plot and another for the sub-plot. In each replication, main plot treatments are first randomly assigned to the main plots followed by a random assignment of the sub-plot treatments within each main plot. This procedure is followed for all replications. A possible layout of a split plot experiment with four main plot treatments (a=4), three sub-plot treatments (b=3), and four replications (r=4) is given below:

Rep. I

b_1	b_3	b_2	b_2
b_3	b_2	b_1	b_3
b_2	b_1	b_3	b_1
a_4	a_2	a_1	a_3

Rep. II

b_3	b_1	b_2	b_1
b_1	b_2	b_1	b_3
b_2	b_3	b_3	b_2
a_1	a_4	a_2	a_3

Rep. III

b_3	b_1	b_2	b_3
b_2	b_3	b_3	b_2
b_1	b_2	b_1	b_1
a_3	a_2	a_4	a_1

Rep. IV

b_2	b_3	b_3	b_1
b_1	b_2	b_1	b_2
b_3	b_1	b_2	b_3
a_1	a_4	a_3	a_2

The above layout has the following important features –

- The size of the main plot is *b* times the size of the sub-plot,
- Each main plot treatment is tested *r* times whereas each sub-plot treatment is tested *ar* times, thus the number of times a sub-plot treatment is tested will always be larger than that for the main plot and is the primary reason for more precision for the sub-plot treatments relative to the main plot treatments.

This concept of splitting each plot may be extended further to accommodate the application of additional factors. An extension of this design is called the split-split plot design where the sub-plot is further divided to include a third factor in the experiment. The design allows for 3 different levels of precision associated with the 3 factors. That is, the degree of precision associated with the main factor is lowest, while the degree of precision associated with the sub-sub plot is the highest.

1.3 Model

The model for simple split plot design is

$$Y_{ijk} = \mu + \rho_i + \tau_j + \delta_{ij} + \beta_k + (\tau\beta)_{jk} + \varepsilon_{ijk} \quad \forall\, i = 1,2,\ldots,r,\ j = 1,2,\ldots,a,\ k = 1,2,\ldots b,$$

where,

Y_{ijk} : observation corresponding to k^{th} level of sub-plot factor(B), j^{th} level of main plot factor(A) and the i^{th} replication.

μ : general mean

ρ_i : i^{th} block effect

τ_j : j^{th} main plot treatment effect

β_k : k^{th} sub-plot treatment effect

$(\tau\beta)_{jk}$: interaction between j^{th} level of main-plot treatment and the k^{th} level of sub-plot treatment

The error components δ_{ij} and ε_{ijk} are independently and normally distributed with means zero and respective variances σ^2_δ and σ^2_ε.

1.4. Analysis

Whole-Plot analysis:

This part of the analysis is based on comparisons of whole-plot totals:

- The levels of A are assigned to the whole plots within blocks according to a randomized complete block design, and so the sum of squares for A needs no block adjustment. There are a –1 degrees of freedom for A, so the sum of squares is given by

  $$ssA = \sum_j y^2_{.j.} / rb - y^2_{...} / rab$$

  [The "dot" notation means "add over all values of the subscript replaced with a dot"]

- There are r –1 degrees of freedom for blocks, giving a block sum of squares of

  $$ssR = \sum_i y^2_{i..} / ab - y^2_{...} / rab$$

- There are a whole plots nested within each of the r blocks, so there are, in total, $r(a-1)$ whole-plot degrees of freedom. Of these, a –1 are used to measure the effects of A leaving $(r-1)(a-1)$ degrees of freedom for whole-plot error. Equivalently, this can be obtained by the subtraction of the block and A degrees of freedom from the whole-plot total degrees of freedom i.e. $(ra-1)-(r-1)-(a-1) = (r-1)(a-1)$.

  So, the whole plot error sum of squares, is obtained as

  $$ssE_1 = \sum_i \sum_j y^2_{ij.} / b - y^2_{...} / rab - ssR - ssA$$

- The whole plot error mean square $msE_1 = ssE_1 / (r-1)(a-1)$, is used as the error estimate to test the significance of whole plot factor(A).

Sub-plot analysis:

This part of the analysis is based on the observations arising from the split-plots within whole plots:

- There are rab –1 total degrees of freedom, and the total sum of squares is

$$sstot = \sum_i \sum_j \sum_k y_{ijk}^2 - y_{...}^2 / rab$$

- Due to the fact that all levels of B are observed in every whole plot as in a randomized complete block design, the sum of squares for B needs no adjustment for whole plots, and is given by -

  $ssB = \sum_k y_{..k}^2 / ra - y_{...}^2 / rab$, corresponding to b –1 degrees of freedom.

- The interaction between the factors A and B is also calculated as part of the split-plot analysis. Again, due to the complete block structure of both the whole-plot design and the split-plot design, the interaction sum of squares needs no adjustment for blocks. The number of interaction degrees of freedom is $(a-1)(b-1)$, and the sum of squares is

  $ss(AB) = \sum_j \sum_k y_{.jk}^2 / r - y_{...}^2 / rab - ssA - ssB$

- Since there are b split plots nested within the ra whole plots, there are, in total, $ra(b-1)$ split-plot degrees of freedom. Of these, b –1 are used to measure the main effect of B, and $(a-1)(b-1)$ are used to measure the AB interaction, leaving $ra(b-1) - (b-1) - (a-1)(b-1) = a(r-1)(b-1)$ degrees of freedom for error. Equivalently, this can be obtained by subtraction of the whole plot, B, and AB degrees of freedom from the total i.e. $(rab-1) - (ra-1) - (b-1) - (a-1)(b-1) = a(r-1)(b-1)$.

  The split-plot error sum of squares can be calculated by subtraction:
  $ssE_2 = sstot - ssR - ssA - ssE_1 - ssB - ss(AB)$.

- The split-plot error mean square $msE_2 = ssE_2 / a(r-1)(b-1)$ is used as the error estimate in testing the significance of split-plot factor(B) and interaction(AB).

- The analysis of variance table is outlined as follows:

ANOVA

Source of Variation	**Degrees of Freedom**	**Sum of Squares**	**Mean Square**	**F**
Whole plot analysis				
Replication	r-1	ssR	-	-
Main plot treatment(A)	a-1	ssA	MsA	msA/msE_1
Main plot error(E_1)	(r-1)(a-1)	ssE_1	$msE_1 = E_a$	
Sub-plot analysis				
Sub-plot treatment(B)	b-1	ssB	MsB	msB/msE_2
Interaction (AxB)	(a-1)(b-1)	ss(AB)	ms(AB)	$ms(AB)/msE_2$
Sub-plot error(E_2)	a(r-1)(b-1)	ssE_2	$msE_2 = E_b$	
Total	rab-1	sstot		

1.5. Standard Errors and Critical Differences:

Estimate of S.E. of difference between two main plot treatment means = $\sqrt{\frac{2E_a}{rb}}$

Estimate of S.E. of difference between two sub-plot treatment means = $\sqrt{\frac{2E_b}{ra}}$

Estimate of S.E. of difference between two sub-plot treatment means at the same level of main plot treatment = $\sqrt{\frac{2E_b}{r}}$

Estimate of S.E. of difference between two main plot treatment means at the same or different levels of sub-plot treatment = $\sqrt{\frac{2[(b-1)E_b + E_a]}{rb}}$

Critical difference is obtained by multiplying the $S.E_d$ by $t_{5\%}$ table value for respective error d.f. for (i), (ii) & (iii). For (iv), as the standard error of mean difference involves two error terms, we use the following equation to compute the weighted t values:

$$t = \frac{(b-1)E_b t_b + E_a t_a}{(b-1)E_b + E_a}$$

where t_a and t_b are t-values at error d.f. (E_a) and error d.f.(E_b) respectively.

Example: In a study carried by agronomists to determine if major differences in yield response to N fertilization exist among different varieties of jowar, the main plot treatments were three varieties of jowar (V_1: CO-18, V_2: CO-19 and V_3: C0-22), and the sub-plot treatments were N rates of 0, 30, and 60 Kg/ha. The study was replicated four times, and the data gathered for the experiment are shown in Table 1.

Table 1: Replication-wise yield data.

		N rate, Kg/ha		
Replication	Variety	0	30	60
		Yield, kg per plot		
I	V_1	15.5	17.5	20.8
	V_2	20.5	24.5	30.2
	V_3	15.6	18.2	18.5
II	V_1	18.9	20.2	24.5
	V_2	15.0	20.5	18.9
	V_3	16.0	15.8	18.3
III	V_1	12.9	14.5	13.5
	V_2	20.2	18.5	25.4
	V_3	15.9	20.5	22.5
IV	V_1	12.9	13.5	18.5
	V_2	13.5	17.5	14.9
	V_3	12.5	11.9	10.5

Analyze the data and draw conclusions.

Steps of analysis: Calculate the replication totals(R), and the grand total(G) by first constructing a table for the replication × variety totals shown in Table 1.1, and then a second table for the variety × nitrogen totals as shown in Table 1.2.

Table 1.1 Replication × variety (RA) - table of yield totals.

	Variety			
Replication	V_1	V_2	V_3	Rep.Total(R)
I	53.8	75.2	52.3	181.3
II	63.6	54.4	50.1	168.1
III	40.9	64.1	58.9	163.9
IV	44.9	45.9	34.9	125.7
Variety Total(A)	203.2	239.6	196.2	
Grand Total(G)				639.0

Table 1.2: Variety × Nitrogen (AB) - table of yield totals.

	Variety			
Nitrogen	V_1	V_2	V_3	Nitrogen Total(B)
N_0	60.2	69.2	60.0	189.4
N_1	65.7	81.0	66.4	213.1
N_2	77.3	89.4	69.8	236.5

- Compute the various sums of squares for the main plot analysis by first computing the correction factor:

C.F. $= \frac{G^2}{rab} = \frac{(639)^2}{4\times 3\times 3} = 11342.25$

Total S.S. (sstot) $= [\,(15.5)^2 + (20.5)^2 + \ldots + (10.5)^2\,] - $ C.F.

$= 637.97$

Replication S.S. (ssR) $= \frac{\sum R^2}{ab} - $ C.F.

$= \frac{(181.3)^2 + (168.1)^2 + (163.9)^2 + (125.7)^2}{3\times 3} - 11342.25$

$= 190.08$

S.S. due to Variety (ssA) $= \frac{\sum A^2}{rb} - $ C.F.

$= \frac{(203.2)^2 + (239.6)^2 + (196.2)^2}{4\times 3} - 11342.25$

$= 90.487$

Main plot error S.S. (ssE_1) = $\frac{\sum(RA)^2}{b} - C.F. - ssR - ssA$

$= \frac{(53.8)^2 + (63.6)^2 + ... + (34.9)^2}{3} - 11342.25 - 190.08 - 90.487$

= 174.103

- Compute the various sums of squares for sub-plot analysis:

S.S. due to Nitrogen (ssB) = $\frac{\sum B^2}{ra} - C.F.$

$= \frac{(189.4)^2 + (213.1)^2 + (236.5)^2}{4 \times 3} - 11342.25$

= 92.435

S.S. due to Interaction (A × B) = $\frac{\sum(AB)^2}{r} - C.F. - ssA - ssB$

$= \frac{(60.2)^2 + (65.7)^2 + ... + (69.8)^2}{4} - 11342.25 - 90.487 - 92.435$

= 9.533

Sub-plot error S.S. (ssE_2) = Total S.S. − All other sum of squares

= 637.97 – (190.08 + 90.487 + 174.103 + 92.435 +9.533)

= 81.332

- Calculate the mean square for each source of variation by dividing the S.S. by its corresponding degrees of freedom and compute the F value for each effect that needs to be tested, by dividing each mean square by the corresponding error mean square, as shown in Table 1.3.

Table 1.3 ANOVA results.

Source of variation	**Degrees of freedom**	**Sum of Squares**	**Mean Square**	**F**
Replication	3	190.08	63.360	
Variety(A)	2	90.487	45.243	1.56^{ns}
Error(a)	6	174.103	29.017(E_a)	
Nitrogen(B)	2	92.435	46.218	10.23^{**}
Variety×Nitrogen (A×B)	4	9.533	2.383	<1
Error(b)	18	81.332	4.518 (E_b)	
Total	35	637.97		

^{ns} – not significant, ** - significant at 1% level.

- Compute the coefficient of variation for the main plot and sub-plot as:

  $cv(a) = \frac{\sqrt{E_a}}{G.M.} \times 100$, and $cv(b) = \frac{\sqrt{E_b}}{G.M.} \times 100$ respectively.

- Compute standard errors and to make specific comparisons among treatment means compute respective critical differences only when F-tests show significance differences and interpret.

- Conclusion: There was no significant difference among variety means. Yield was significantly affected by nitrogen. However, the interaction between N rate and variety was not significant. All the varieties gave significant response to 30 kg N/ha as well as to 60 kg N/ha.

2. Strip Plot Design

2.1 Introduction

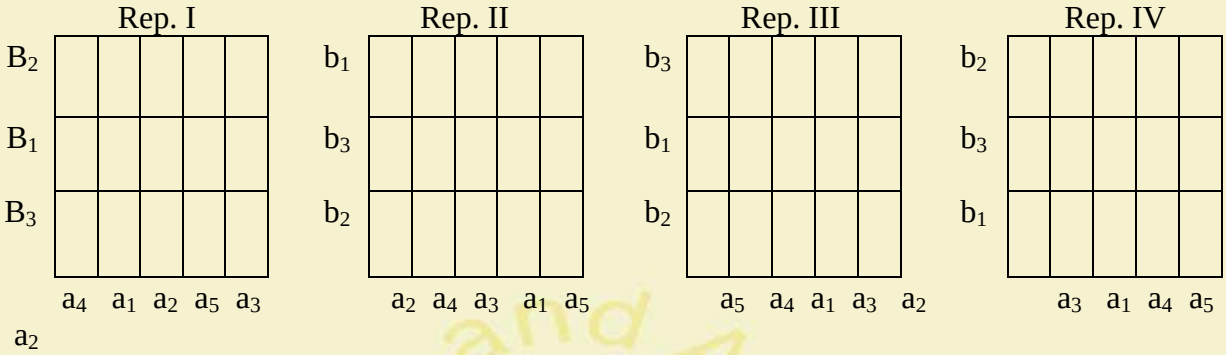
Sometimes situation arises when two factors each requiring larger experimental units are to be tested in the same experiment, e.g., suppose four levels of spacing and three levels of methods of ploughing are to be tested in the same experiment. Here both the factors require large experimental units. If the combination of the two factors at all possible levels are allotted in a R.B.D. in the normal way, the experimental plots shall have to be very large thereby bringing heterogeneity. So, it will not be appropriate. On the other hand if one factor(spacing) is taken in main plots and other factor(methods of ploughing) is taken in sub-plots within main plots, the sub-plots shall have to be large enough. Hence split plot design also will not be appropriate. In such situations a design called Strip plot design is adopted.

The strip plot is a 2-factor design that allows for greater precision in the measurement of the interaction effect while sacrificing the degree of precision on the main effects. The experimental area is divided into three plots, namely the vertical-strip plot, the horizontal-strip plot, and the intersection plot. We allocate factors A and B, respectively, to the vertical and horizontal-strip plots, and allow the intersection plot to accommodate the interaction between these two factors. As in the split plot design, the vertical and the horizontal plots are perpendicular to each other. However, in the strip plot design the relationship between the vertical and horizontal plot sizes is not as distinct as the main and sub-plots were in the split plot design. The sub-plot treatments instead of being randomized independently within each main plot as in the case of split plot design are arranged in strips across each replication. The intersection plot, which is one of the characteristics of the design, is the smallest in size.

2.2. Randomization and Layout:

In this design each block is divided into number of vertical and horizontal strips depending on the levels of the respective factors. Let A represent the vertical factor with *a* levels, B represent the horizontal factor with *b* levels and *r* represent the number of replications. To layout the experiment, the experimental area is divided into *r* blocks. Each block is divided into *b* horizontal strips and *b* treatments are randomly assigned to these strips in each of the *r* blocks separately and independently. Then each block is divided into *a* vertical strips and *a* treatments are randomly assigned to these strips in each of the *r* blocks separately and

independently. A possible layout of a strip plot experiment with a =5 (a_1, a_2, a_3, a_4, and a_5) , b =3 (b_1, b_2, and b_3) and four replications is given below:



The strip plot design sacrifices precision on the main effects of both the factors in order to provide higher precision on the interaction which will generally be more accurately determined than in either randomized blocks or simple split plot design. Consequently this design is not recommended unless practical considerations necessitate its use or unless the interaction is the principle object of study.

2.3. Model

The model for strip plot design is

$Y_{ijk} = \mu + \rho_i + \alpha_j + (\rho\alpha)_{ij} + \beta_k + (\rho\beta)_{ik} + (\alpha\beta)_{jk} + \varepsilon_{ijk}$ $\forall$ *i = 1,2, …,r, j = 1,2, …,a, k = 1,2, …b*

where,

Y_{ijk} : observation corresponding to k^{th} level of factor A, j^{th} level of factor B and i^{th} replication.

μ : general mean

ρ_i : i^{th} block effect

α_j : effect of j^{th} level of factor A

β_k : effect of k^{th} level of factor B

$(\alpha\beta)_{jk}$: interaction between j^{th} level of factor A and the k^{th} level of factor B

The error components $(\rho\alpha)_{ij}$, $(\rho\beta)_{ik}$ and ε_{ijk} are independently and normally distributed with means zero and respective variances σ^2_a, σ^2_b, and σ^2_ε.

2.4. Analysis

In statistical analysis separate estimates of error are obtained for main effects of the factor, A and B and for their interaction AB. Thus there will be three error mean squares applicable for testing the significance of main effects of the factors and their interaction separately.

Suppose 4 levels of spacings(A) and 3 levels of methods(B) of ploughing are to be tested in the same experiment. Each replication is divided into 4 strips vertically and into 3 strips horizontally. In the vertical strips the four different levels of spacings are allotted randomly and in the horizontal strips three methods of ploughing are allotted randomly. Let there be 4

replications(R). The analysis of variance is carried out in three parts viz. vertical strip analysis, horizontal strip analysis and interaction analysis as follows:

- Form spacing × replication (A × R) table of yield totals and from this table compute the S.S. due to replication, S.S. due to spacings and S.S. due to interaction - Replication × Spacing i.e. error(a).
- Form method × replication (B × R) table of yield totals and from this table compute the S.S. due to methods and S.S. due to interaction - Replication × Method i.e. error(b).
- Form spacing × method (A × B) table of yield totals and from this table compute the S.S. due to interaction - Spacing × Method.
- Total S.S. will be obtained as usual by considering all the observations of the experiment and the error S.S. i.e. error(c) will be obtained by subtracting from total S.S. all the S.S. for various sources.
- Now, calculate the mean square for each source of variation by dividing each sum of squares by its respective degrees of freedom.
- Compute the F-value for each source of variation by dividing each mean square by the corresponding error term.
- The analysis of variance table is outlined as follows:

ANOVA

Source of Variation	**Degrees of Freedom**	**Sum of Squares**	**Mean Square**	**F**
Replication(R)	(r-1)= 3	ssR	-	-
Spacing(A)	(a-1)= 3	ssA	MsA	msA/msE_1
Error(a)	(r-1)(a-1)= 9	ssE_1	$msE_1 = E_a$	
Method(B)	(b-1)= 2	ssB	MsB	msB/msE_2
Error(b)	(r-1)(b-1)= 6	ssE_2	$msE_2 = E_b$	
Spacing×Method (A×B)	(a-1)(b-1)= 6	ss(AB)	ms(AB)	$ms(AB)/msE_3$
Error(c)	(r-1)(a-1)(b-1)=18	ssE_3	$msE_3 = E_c$	
Total	(rab-1)= 47	sstot		

2.5 Standard Errors and Critical Differences:

Estimate of S.E. of difference between two A level means = $\sqrt{\frac{2E_a}{rb}}$

Estimate of S.E. of difference between two B level means = $\sqrt{\frac{2E_b}{ra}}$

Estimate of S.E. of difference between two A level means at the same level of B means = $\sqrt{\frac{2[(b-1)E_c + E_a]}{rb}}$

Estimate of S.E. of difference between two B level means at the same level of A means = $\sqrt{\frac{2[(a-1)E_c + E_b]}{ra}}$

Critical difference is obtained by multiplying the $S.E_d$ by $t_{5\%}$ table value for respective error d.f. for (i) & (ii). For (iii) & (iv), as the standard error of mean difference involves two error terms, we use the following equation to compute the weighted t values:

$$t = \frac{(b-1)E_c t_c + E_a t_a}{(b-1)E_c + E_a}, \text{ and } t = \frac{(a-1)E_c t_c + E_b t_b}{(a-1)E_c + E_b} \text{ respectively,}$$

where t_a, t_b, and t_c are t-values at error d.f. (E_a), error d.f.(E_b) and error d.f.(E_c) respectively.

SAS input statements for the split plot experiment

```
data split plot;
input rep var nit yield;
cards;
 1  1  0  15.5
 1  1  1  17.5
 1  1  2  20.8
 1  2  0  20.5
 1  2  1  24.5
 1  2  2  30.2
 1  3  0  15.6
 1  3  1  18.2
 1  3  2  18.5
 2  1  0  18.9
 2  1  1  20.2
 2  1  2  24.5
 2  2  0  15.0
 2  2  1  20.5
 2  2  2  18.9
 2  3  0  16.0
 2  3  1  15.8
 2  3  2  18.3
 3  1  0  12.9
 3  1  1  14.5
 3  1  2  13.5
 3  2  0  20.2
 3  2  1  18.5
 3  2  2  25.4
 3  3  0  15.9
 3  3  1  20.5
 3  3  2  22.5
 4  1  0  12.9
 4  1  1  13.5
 4  1  2  18.5
 4  2  0  13.5
```

```
 4  2  1  17.5
 4  2  2  14.9
 4  3  0  12.5
 4  3  1  11.9
 4  3  2  10.5
;
proc print;
proc glm;
class rep var nit;
model yield = rep var rep* var nit var*nit;
test h = var e = rep*var;
mean var nit var*nit;
run;
```

Chapter 12
Analysis of Covariance

Any scientific experiment is performed to know something that is unknown about a group of treatments and to test certain hypothesis about the corresponding treatment effect.

When variability of experimental units is small relative to the treatment differences and the experimenter do not wishes to use experimental design, then just take large number of observations on each treatment effect and compute its mean. The variation around mean can be made as small as desired by taking more observations.

When there is considerable variation among observations on the same treatment and it is not possible to take an unlimited number of observations, the techniques used for reducing the variation are

(i) use of proper experimental design and

(ii) use of concomitant variables.

The use of concomitant variables is accomplished through the technique of analysis of covariance. If both the techniques fail to control the experimental variability then the number of replications of different treatments (in other words, the number of experimental units) are needed to be increased to a point where adequate control of variability is attained.

Introduction to analysis of covariance model

In the linear model

$$Y = X_1\beta_1 + X_2\beta_2 + ... + X_p\beta_p + \varepsilon,$$

if the explanatory variables are quantitative variables as well as indicator variables, i.e., some of them are qualitative and some are quantitative, then the linear model is termed as analysis of covariance (ANCOVA) model.

Note that the indicator variables do not provide as much information as the quantitative variables. For example, the quantitative observations on age can be converted into indicator variable. Let an indictor variable be

$$D = \begin{cases} 1 & \text{if age} \geq 17 \text{ years} \\ 0 & \text{if age} < 17 \text{ years.} \end{cases}$$

Now the following quantitative values of age can be changed into indicator variables.

Ages (years)	Ages
14	0
15	0
16	0
17	1
20	1
21	1
22	1

In many real application, some variables may be quantitative and others may be qualitative. In such cases, ANCOVA provides a way out.

It helps is reducing the sum of squares due to error which in turn reflects the better model adequacy diagnostics.

See how does this work:

In one way model: $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, we have $TSS_1 = SSA_1 + SSE_1$

In two way model: $Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$, we have $TSS_2 = SSA_2 + SSB_2 + SSE_2$

In three way model: $Y_{ij} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ik}$, we have $TSS_3 = SSA_3 + SSB_3 + SS\gamma_3 + SSE_3$

If we have a given data set, then ideally

$$TSS_1 = TSS_2 = TSS_3$$
$$SSA_1 = SSA_2 = SSA_3;$$
$$SSB_2 = SSB_3$$

So $\text{SSE}_1 \geq SSE_2 \geq SSE_3$.

Note that in the construction of F-statistics, we use $\dfrac{SS(effects)/df}{SSE/df}$.

So F-statistic essentially depends on the $SSEs$.

Smaller $SSE \Rightarrow$ large $F \Rightarrow$ more chance of rejection.

Since SSA, SSB etc. here are based on dummy variables, so obviously if SSA, SSB, etc. are based on quantitative variables, they will provide more information. Such ideas are used in ANCOVA models and we construct the model by incorporating the quantitative explanatory variables in ANOVA models.

In another example, suppose our interest is to compare several different kinds of feed for their ability to put weight on animals. If we use ANOVA, then we use the final weights at the end of experiment. However, final weights of the animals depend upon the initial weight of the animals at the beginning of the experiment as well as upon the difference in feeds.

Use of ANCOVA models enables us to adjust or correct these initial differences.

ANCOVA is useful for improving the precision of an experiment. Suppose response Y is linearly related to covariate X (or **concomitant variable**). Suppose experimenter cannot control X but can observe it. ANCOVA involves adjusting Y for the effect of X. If such an adjustment is not made, then the X can inflate the error mean square and makes the true differences is Y due to treatment harder to detect.

If, for a given experimental material, the use of proper experimental design cannot control the experimental variation, the use of concomitant variables (which are related to experimental material) may be effective in reducing the variability.
Consider the one way classification model as

$$E(Y_{ij} = \beta_i \qquad i = 1,..., p,\ j = 1,..., N_i,$$
$$Var(Y_{ij}) = \sigma^2.$$

If usual analysis of variance for testing the hypothesis of **equality of treatment effects** shows a highly significant difference in the treatment effects due to some factors affecting the experiment, then consider the model which takes into account this effect

$$E(Y_{ij}) = \beta_i + \gamma t_{ij} \qquad i = 1,..., p,\ j = 1,..., N_i,$$
$$Var(Y_{ij}) = \sigma^2$$

where t_{ij} are the observations on concomitant variables (which are related to X_{ij}) and γ is the regression coefficient associated with t_{ij}. With this model, the variability of treatment effects can be considerably reduced.

For example, in any agricultural experimental, if the experimental units are plots of land then, t_{ij} can be measure of fertility characteristic of the j^{th} plot receiving i^{th} treatment and X_{ij} can be yield.

In another example, if experimental units are animals and suppose the objective is to compare the growth rates of groups of animals receiving different diets. Note that the observed differences in growth rates can be attributed to diet only if all the animals are similar in some observable characteristics like weight, age etc. which influence the growth rates.

In the absence of similarity, use t_{ij} which is the weight or age of j^{th} animal receiving i^{th} treatment.

If we consider the quadratic regression in t_{ij} then in

$$E(Y_{ij}) = \beta_i + \gamma_i t_{ij} + \gamma_2 t_{ij}^2, \quad i = 1,..., p, j = 1,..., n_i,$$
$$Var(Y_{ij}) = \sigma^2.$$

ANCOVA in this case is the same as ANCOVA with two concomitant variables t_{ij} and t_{ij}^2.

In two way classification with one observation per cell,

$E(Y_{ij}) = \mu + \alpha_i + \beta_j + \gamma t_{ij}, \quad i = 1,..., I, \ j = 1,..., J$

or

$E(Y_{ij}) = \mu + \alpha_i + \beta_j + \gamma_i t_{ij} + \gamma_2 w_{ij}$

with $\sum_i \alpha_i = 0, \ \sum_j \beta_j = 0,$

then (y_{ij}, t_{ij}) or (y_{ij}, t_{ij}, w_{ij}) are the observations in $(i, j)^{th}$ cell and t_{ij}, w_{ij} are the concomitment variables.

The concomitant variables can be fixed on random.

We consider the case of fixed concomitant variables only.

One way classification

Let Y_{ij} $(j = 1 \ldots n_i,\ i = 1 \ldots p)$ be a random sample of size n_i from i^{th} normal populations with mean

$$\mu_{ij} = E(Y_{ij}) = \beta_i + \gamma t_{ij}$$
$$Var(Y_{ij}) = \sigma^2$$

where β_i, γ and σ^2 are the unknown parameters, t_{ij} are known constants which are the observations on a concomitant variable.

The null hypothesis is

$$H_0 : \beta_1 = \ldots = \beta_p .$$

Let

$$\bar{y}_{io} = \frac{1}{n_i}\sum_j y_{ij};\ \ \bar{y}_{oj} = \frac{1}{p}\sum_i y_{ij},\ \ \bar{y}_{oo} = \frac{1}{n}\sum_i\sum_j y_{ij}$$
$$\bar{t}_{io} = \frac{1}{n_i}\sum_j t_{ij};\ \ \bar{t}_{oj} = \frac{1}{p}\sum_i t_{ij},\ \ \bar{t}_{oo} = \frac{1}{n}\sum_i\sum_j t_{ij}$$
$$n = \sum_i n_i .$$

Under the whole parametric space (π_Ω), use likelihood ratio test for which we obtain the $\hat{\beta}_i$'s and $\hat{\gamma}$ using the least squares principle or maximum likelihood estimation as follows:

Minimize $S = \sum_i\sum_j (y_{ij} - \mu_{ij})^2$

$$= \sum_i\sum_j (y_{ij} - \beta_i - \gamma t_{ij})^2$$

$\frac{\partial S}{\partial \beta_i} = 0$ for fixed γ

$\Rightarrow \beta_i = \bar{y}_{io} - \gamma \bar{t}_{io}$

Put β_i in S and minimize the function by $\frac{\partial S}{\partial \gamma} = 0$,

i.e., minimize $\sum_i\sum_j \left[y_{ij} - \bar{y}_{io} - \gamma(t_{ij} - \bar{t}_{io}) \right]^2$ with respect to γ gives

$$\hat{\gamma} = \frac{\sum_i\sum_j (y_{ij} - \bar{y}_{io})(t_{ij} - \bar{t}_{io})}{\sum_i\sum_j (t_{ij} - \bar{t}_{io})^2}.$$

Thus $\hat{\beta}_i = \bar{y}_{io} - \hat{\gamma}\bar{t}_{io}$

$$\hat{\mu}_{ij} = \hat{\beta}_i + \hat{\gamma} t_{ij}.$$

Since $y_{ij} - \hat{\mu}_{ij} = y_{ij} - \hat{\beta}_i - \hat{\gamma} t_{ij}$

$= y_{ij} - \overline{y}_{io} - \hat{\gamma}(t_{ij} - \overline{t}_{io}),$

we have

$$\sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2 = \sum\sum (y_{ij} - \overline{y}_{io})^2 - \frac{\left[\sum_i \sum_j (y_{ij} - \overline{y}_{io})(t_{ij} - \overline{t}_{io})\right]}{\sum_i \sum_j (t_{ij} - \overline{t}_{io})^2}.$$

Under $H_0 : \beta_1 = ... = \beta_p = \beta$ (say), consider $S_w = \sum_i \sum_j \left[y_{ij} - \beta - \gamma t_{ij} \right]^2$ and minimize S_w under sample space (π_w),

$$\frac{\partial S_w}{\partial \beta} = 0,$$

$$\frac{\partial S_w}{\partial \gamma} = 0$$

$$\Rightarrow \hat{\hat{\beta}} = \overline{y}_{oo} - \hat{\hat{\gamma}} \overline{t}_{oo}$$

$$\hat{\hat{\gamma}} = \frac{\sum_i \sum_j (y_{ij} - \overline{y}_{oo})(t_{ij} - \overline{t}_{oo})}{\sum_i \sum_j (t_{ij} - \overline{t}_{oo})^2}$$

$$\hat{\hat{\mu}}_{ij} = \hat{\hat{\beta}} + \hat{\hat{\gamma}} t_{ij}.$$

Hence

$$\sum_i \sum_j (y_{ij} - \hat{\hat{\mu}}_{ij})^2 = \sum_i \sum_j (y_{ij} - \overline{y}_{oo})^2 - \frac{\left[\sum_i \sum_j (y_{ij} - \overline{y}_{oo})(t_{ij} - \overline{t}_{oo})\right]^2}{\sum_i \sum_j (t_{ij} - \overline{t}_{oo})^2}$$

and

$$\sum_i \sum_j (\hat{\mu}_{ij} - \hat{\hat{\mu}}_{ij})^2 = \sum_i \sum_j \left[(\overline{y}_i - \overline{y}_{oo}) + \hat{\gamma}(t_{ij} - \overline{t}_{io}) - \hat{\hat{\gamma}}(t_{ij} - \overline{t}_{oo}) \right]^2.$$

The likelihood ratio test statistic in this case is given by

$$\lambda = \frac{\max_{w} L(\beta,\gamma,\sigma^2)}{\max_{\Omega} L(\beta,\gamma,\sigma^2)}$$
$$= \frac{\sum_i \sum_j (\hat{\mu}_{ij} - \hat{\hat{\mu}}_{ij})^2}{\sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2}.$$

Now we use the following theorems:

Theorem 1: Let $Y = (Y_1, Y_2, ..., Y_n)'$ follow a multivariate normal distribution $N(\mu, \Sigma)$ with mean vector μ and positive definite covariance matrix Σ. Then $Y'AY$ follows a noncentral chi-square distribution with p degrees of freedom and noncentrality parameter $\mu' A\mu$, i.e., $\chi^2(p, \mu' A\mu)$ if and only if ΣA is an idempotent matrix of rank p.

Theorem 2: Let $Y = (Y_1, Y_2, ..., Y_n)'$ follows a multivariate normal distribution $N(\mu, \Sigma)$ with mean vector μ and positive definite covariance matrix Σ. Let $Y'A_1Y$ follows $\chi^2(p_1, \mu' A_1\mu)$ and $Y'A_2Y$ follows $\chi^2(p_2, \mu' A_2\mu)$. Then $Y'A_1Y$ and $Y'A_2Y$ are independently distributed if $A_1\Sigma A_2 = 0$.

Theorem 3: Let $Y = (Y_1, Y_2, ..., Y_n)'$ follows a multivariate normal distribution $N(\mu, \sigma^2 I)$, then the maximum likelihood (or least squares) estimator $L'\hat{\beta}$ of estimable linear parametric function is independently distributed of $\hat{\sigma}^2; L\hat{\beta}$ follow $N\left[L'\beta, L'(X'X)^{-1}L\right]$ and $\frac{n\hat{\sigma}^2}{\sigma^2}$ follows $\chi^2(n-p)$ where $rank(X) = p$.

Using these theorems on the independence of quadratic forms and dividing the numerator and denominator by respective degrees of freedom, we have

$$F = \frac{n-p-1}{p-1} \frac{\sum_i \sum_j (\hat{\mu}_{ij} - \hat{\hat{\mu}}_{ij})^2}{\sum \sum (y_{ij} - \hat{\mu}_{ij})^2} \sim F(p-1, n-p) \text{ under } H_0$$

So reject H_0 whenever $F \geq F_{1-\alpha}(p-1, n-p)$ at α level of significance.

The terms involved in λ can be simplified for computational convenience follows:

We can write

$$\sum_i \sum_j (y_{ij} - \hat{\hat{\mu}}_{ij})^2$$

$$= \sum_i \sum_j \left[y_{ij} - \hat{\hat{\beta}} - \hat{\hat{\gamma}} t_{ij} \right]^2$$

$$= \sum_i \sum_j \left[(y_{ij} - \bar{y}_{oo}) - \hat{\hat{\gamma}} (t_{ij} - \bar{t}_{oo}) \right]^2$$

$$= \sum_i \sum_j \left[(y_{ij} - \bar{y}_{oo}) - \hat{\hat{\gamma}} (t_{ij} - \bar{t}_{oo}) + \hat{\gamma} (t_{ij} - \bar{t}_{io}) - \hat{\gamma} (t_{ij} - \hat{t}_{io}) \right]^2$$

$$= \sum_i \sum_j \left[(y_{ij} - \bar{y}_{io}) - \hat{\gamma} (t_{ij} - \hat{t}_{io}) \right]^2$$

$$= \sum_i \sum_j \left[(y_{ij} - \bar{y}_{oo}) + \hat{\gamma} (t_{ij} - \bar{t}_{io}) - \hat{\hat{\gamma}} (t_{ij} - \bar{t}_{oo}) \right]^2$$

$$= \sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2 + \sum_i \sum_j (\hat{\mu}_{ij} - \hat{\hat{\mu}}_{ij}).$$

For computational convenience

$$\lambda = \frac{\sum_i \sum_j (\hat{\mu}_{ij} - \hat{\hat{\mu}}_{ij})^2}{\sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2} = \frac{\left(T_{yy} - \frac{T_{yt}^2}{T_{tt}} \right) - \left(E_{yy} - \frac{E_{yt}^2}{E_{tt}} \right)}{\left(E_{yy} - \frac{E_{yt}^2}{E_{yy}} \right)}$$

where

$$T_{yy} = \sum_i \sum_j (y_{ij} - \bar{y}_{oo})^2$$

$$T_{tt} = \sum_i \sum_j (t_{ij} - \bar{t}_{oo})^2$$

$$T_{yt} = \sum_i \sum_j (y_{ij} - \bar{y}_{oo})(t_{ij} - \bar{t}_{oo})$$

$$E_{yy} = \sum_i \sum_j (y_{ij} - \bar{y}_{io})^2$$

$$E_{tt} = \sum_i \sum_j (t_{ij} - \bar{t}_{io})^2$$

$$E_{yt} = \sum_i \sum_j (y_{ij} - \bar{y}_{io})(t_{ij} - \bar{t}_{io}).$$

Analysis of covariance table for one way classification is as follows:

Source of variation	Degrees of freedom	Sum of products			Adjusted sum of squares		F
		yy	yt	tt	Degress of feedom	Sum of squares	
Population	$p-1$	$P_{yy}(=T_{yy}-E_{yy})$	$P_{yt}(=T_{yt}-E_{yt})$	$P_{tt}(=T_{tt}-E_{tt})$	$p-1$	$q_1=q_0-q_2$	$\frac{n-p-1}{p-1}\frac{q_1}{q_2}$
Error	$n-p$	E_{yy}	E_{yt}	E_{tt}	$n-p-1$	$q_2=E_{yy}-\frac{E_{yt}^2}{E_{yy}}$	
Total	$n-1$	T_{yy}	T_{yt}	T_{tt}	$n-2$	$q_0=T_{yy}-\frac{T_{yt}^2}{T_{tt}}$	

If H_0 is rejected, employ multiple comprises methods to determine which of the contrasts in β_i are responsible for this.

For any estimable linear parametric contrast

$$\varphi=\sum_{i=1}^{p}C_i\beta_i \text{ with } \sum_{i=1}^{p}C_i=0,$$

$$\hat{\varphi}=\sum_{i=1}^{p}C_i\hat{\beta}_i=\sum_{i=1}^{p}C_i\bar{y}_i-\hat{\gamma}\sum_{i=1}^{p}C_{ii}\bar{t}_i$$

$$Var(\hat{\gamma})=\frac{\sigma^2}{\sum_i\sum_j(t_{ij}-\bar{t}_i)^2}$$

$$\Rightarrow Var(\hat{\varphi})=\sigma^2\left[\sum_i\frac{C_i^2}{n_i}+\frac{\left(\sum_i C_i\bar{t}_i\right)^2}{\sum_i\sum_j(t_{ij}-\bar{t}_i)^2}\right].$$

Two way classification (with one observations per cell)

Consider the case of two way classification with one observation per cell.

Let $y_{ij} \sim N(\mu_{ij}, \sigma^2)$ be independently distributed with

$$E(y_{ij}) = \mu + \alpha_i + \beta_j + \gamma t_{ij}, \quad i = 1...I, \; j = 1...J$$
$$V(y_{ij}) = \sigma^2$$

where

μ : Grand mean

α_1 : Effect of i^{th} level of A satisfying $\sum_{i}^{I} \alpha_i = 0$

β_1 : Effect of j^{th} level of B satisfying $\sum_{i}^{J} \beta_j = 0$

t_{ij} : observation (known) on concomitant variable.

The null hypothesis under consideration are

$H_{0\alpha} : \alpha_1 = \alpha_2 = ... = \alpha_I = 0$
$H_{0\beta} : \beta_1 = \beta_2 = ... = \beta_J = 0$

Dimension of whole parametric space $(\pi_\Omega) : I + J$

Dimension of sample space $(\pi_{w\alpha}) : J + 1$ under $H_{0\alpha}$

Dimension of sample space $(\pi_{w\beta}) : I + 1$ under $H_{0\beta}$

with respective alternative hypotheses as

$H_{1\alpha}$: At least one pair of $\alpha's$ is not equal

$H_{1\beta}$: At least one pair of $\beta's$ is not equal.

Consider the estimation of parameters under the whole parametric space (π_Ω).

Find minimum value of $\sum_i \sum_j (y_{ij} - \mu_{ij})^2$ under π_Ω.

To do this, minimize

$$\sum_i \sum_j (y_{ij} - \mu - \alpha_i - \beta_j - \gamma t_{ij})^2 .$$

For fixed γ, which gives on solving the least squares estimates (or the maximum likelihood estimates) of the respective parameters as

$$\mu = \bar{y}_{oo} - \gamma \bar{t}_o$$
$$\alpha_i = \bar{y}_i - \bar{y}_{oo} - \gamma(\bar{t}_{io} - \bar{t}_{oo}) \qquad (1)$$
$$\beta_j = \bar{y}_{oj} - \bar{y}_{oo} - \gamma(\bar{t}_{oj} - \bar{t}_{oo}).$$

Under these values of μ, α_i and β_j, the sum of squares $\sum_i \sum_j (y_{ij} - \mu - \alpha_i - \beta_j - \gamma t_{ij})^2$ reduces to

$$\sum_i \sum_j \left[y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo} + \gamma(t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo}) \right]^2. \qquad (2)$$

Now minimization of (2) with respect to γ gives

$$\hat{\gamma} = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})(t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})}{\sum_{i=1}^{I} \sum_{j=1}^{J} (t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})^2}.$$

Using $\hat{\gamma}$, we get from (1)

$$\hat{\mu} = \bar{y}_{oo} - \hat{\gamma} \bar{t}_{oo}$$
$$\hat{\alpha}_i = (\bar{y}_{io} - \bar{y}_{oo}) - \hat{\gamma}(\bar{t}_{io} - \bar{t}_{oo})$$
$$\hat{\beta}_j = (\bar{y}_{oj} - \bar{y}_{oo}) - \hat{\gamma}(\bar{t}_{oj} - \bar{t}_{oo}).$$

Hence

$$\sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2$$
$$= \sum_i \sum_j (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})^2 - \frac{\left[\sum_i \sum_j (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})(t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo}) \right]}{\sum_i \sum_j (t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})^2}$$
$$= E_{yy} - \frac{E_{yt}^2}{E_{tt}}$$

where

$$E_{yy} = \sum_i \sum_j (y_{ij} - y_{io} - \bar{y}_{oj} + \bar{y}_{oo})^2$$
$$E_{yt} = \sum \sum (y_{ij} - y_{io} - \bar{y}_{oj} + \bar{y}_{oo})(t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})$$
$$E_{tt} = \sum_i \sum_j (t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})^2.$$

Case (i) : Test of $H_{0\alpha}$

Minimize $\sum_i \sum_j (y_{ij} - \mu - \beta_j - \gamma t_{ij})^2$ with respect to μ, β_j and γ gives the least squares estimates) (or the maximum likelihood estimates) of respective parameters as

$$\Rightarrow \hat{\hat{\mu}} = \bar{y}_{oo} - \hat{\hat{\gamma}}\bar{t}_{oo}$$

$$\hat{\hat{\beta}}_j = \bar{y}_{oj} - \bar{y}_{oo} - \hat{\hat{\gamma}}(\bar{t}_{oj} - \bar{t}_{oo})$$

$$\hat{\hat{\gamma}} = \frac{\sum_i \sum_j (y_{ij} - \bar{y}_{oj})(t_{ij} - \bar{t}_{oj})}{\sum_i \sum_j (t_{ij} - \bar{t}_{oj})^2} \qquad (3)$$

$$\hat{\hat{\mu}} = \hat{\hat{\mu}} + \hat{\hat{\beta}}_j + \hat{\hat{\gamma}} t_{ij}.$$

Substituting these estimates in (3) we get

$$\sum_i \sum_j (y_{ij} - \hat{\hat{\mu}}_{ij})^2 = \sum_i \sum_j (y_{ij} - \bar{y}_j)^2 - \frac{\left[\sum_i \sum_j (y_{ij} - \bar{y}_{oj})(t_{ij} - \bar{t}_{oj})\right]}{\sum_i \sum_j (t_{ij} - \bar{t}_{oj})^2}$$

$$= E_{yy} + A_{yy} - \frac{\left[E_{yt} + A_{yt}\right]}{E_{tt} + A_{tt}}$$

where

$$A_{yy} = \sum_i J(\bar{y}_{io} - \bar{y}_{oo})^2$$

$$A_{tt} = \sum_i J(\bar{t}_{io} - \bar{t}_{oo})^2$$

$$A_{yt} = \sum_i J(\bar{y}_{io} - \bar{y}_{oo})(\bar{t}_{io} - \bar{t}_{oo})^2$$

$$E_{yy} = \sum_i \sum_j (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})^2$$

$$E_{tt} = \sum_i \sum_j (t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo})^2$$

$$E_{yt} = \sum_i \sum_j (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})(t_{ij} - \bar{t}_{io} - \bar{t}_{oj} + \bar{t}_{oo}).$$

Thus the likelihood ratio test statistic for testing $H_{0\alpha}$ is

$$\lambda_1 = \frac{\sum_i \sum_j (y_{ij} - \hat{\hat{\mu}}_{ij})^2 - \sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2}{\sum_i \sum_j (y_{ij} - \hat{\mu}_{ij})^2}.$$

Adjusting with degrees of freedom and using the earlier result for the independence of two quadratic forms and their distribution

$$F_1 = \frac{(IJ-I-J)}{(I-1)}\left[\frac{\sum_i\sum_j (y_{ij}-\hat{\tilde{\mu}}_{ij})^2 - \sum_i\sum_j (y_{ij}-\hat{\mu}_{ij})^2}{\sum_i\sum_j (y_{ij}-\hat{\mu}_{ij})^2}\right] \sim F(I-1, IJ-I-J) \text{ under } H_{o\alpha}\ .$$

So the decision rule is to reject $H_{o\alpha}$ whenever $F_1 > F_{1-\alpha}(I-1, IJ-I-J)$.

Case b: Test of $H_{0\beta}$

Minimize $\sum_i\sum_j (y_{ij}-\mu-\alpha_i-\gamma t_{ij})^2$ with respect to μ, α_i and γ gives the least squares estimates (or maximum likelihood estimates) of respective parameters as

$$\begin{aligned}
&\hat{\mu} = \bar{y}_{oo} - \tilde{\gamma}\bar{t}_{oo} \\
&\tilde{\alpha}_j = \bar{y}_{io} - \bar{y}_{oo} - \tilde{\gamma}(\bar{t}_{io} - \bar{t}_{oo}) \\
&\tilde{\gamma} = \frac{\sum_i\sum_j (y_{ij}-\bar{y}_{io})(t_{ij}-\bar{t}_{io})}{\sum_i\sum_j (t_{ij}-\bar{t}_{io})^2} \qquad (4) \\
&\tilde{\mu}_{ij} = \tilde{\mu} + \tilde{\alpha}_i + \tilde{\gamma}_{ij}.
\end{aligned}$$

From (4), we get

$$\begin{aligned}
\sum_i\sum_j (y_{ij}-\tilde{\mu}_{ij})^2 &= \sum_i\sum_j (y_{ij}-\bar{y}_{io})^2 - \frac{\left[\sum_i\sum_j (y_{ij}-\bar{y}_{io})(t_{ij}-\bar{t}_{oj})\right]}{\sum_i\sum_j (t_{ij}-\bar{t}_{io})^2} \\
&= E_{yy} + B_{yy} - \frac{\left[E_{yt}+B_{yt}\right]^2}{B_{tt}}
\end{aligned}$$

$$B_{yy} = \sum_j I(\bar{y}_{oj}-\bar{y}_{oo})^2$$

where $B_{tt} = \sum_j I(\bar{t}_{oj}-\bar{t}_{oo})^2$

$$B_{yt} = \sum_j I(\bar{y}_{io}-\bar{y}_{oo})(\bar{t}_{oj}-\bar{t}_{oo})^2.$$

Thus the likelihood ratio test statistic for testing $H_{0\beta}$ is

$$F_2 = \frac{(IJ-I-J)}{(J-1)}\left[\frac{\sum_i\sum_j (y_{ij}-\tilde{\mu}_{ij})^2 - \sum_i\sum_j (y_{ij}-\hat{\mu}_{ij})^2}{\sum_i\sum_j (y_{ij}-\hat{\mu}_{ij})^2}\right] \sim F(J-1, IJ-I-J) \text{ under } H_{o\beta}.$$

So the decision rule is to reject $H_{0\beta}$ whenever $F_2 \geq F_{1-\alpha}(J-1, IJ-I-J)$.

If $H_{o\alpha}$ is rejected, use multiple comparison methods to determine which of the contrasts α_i are responsible for this rejection. The same is true for $H_{o\beta}$.

The analysis of covariance table for two way classification is as follows:

Source of variation	Degrees of freedom	Sum of products yy	yt	tt		F
Between evels of A	$I-1$	A_{yy}	A_{yt}	A_{tt}	$I-1$ $q_0 = q_3 - q_2$	$F_1 = \frac{IJ-I-J}{I-1}\frac{q_0}{q_2}$
Between levels of B	$J-1$	B_{yy}	B_{yt}	B_{tt}	$J-1$ $q_1 = q_4 - q_2$	$F_2 = \frac{IJ-I-J}{J-1}\frac{q_1}{q_2}$
Error	$(I-1)(J-1)$	E_{yy}	E_{yt}	E_{tt}	$IJ-I-J$ $q_2 = E_{yy} - \frac{E_{yt}^2}{E_{tt}}$	
Total	$IJ-1$	T_{yy}	T_{yt}	T_{tt}	$IJ-2$	
Error + levels of A	$IJ-J$				$q_3 = (A_{yy}+E_{yy}) - \frac{(A_{yt}+E_{yt})^2}{A_{tt}+E_{tt}}$	
Error + levels of B	$IJ-I$				$q_4 = (B_{yy}+E_{yt}) - \frac{(B_{yt}+E_{yt})^2}{B_{tt}+E_{tt}}$	

LATTICE DESIGNS

SONAWANE MAHESH KUMAR NAMDEVRAO
M.Sc. (Agricultural Statistics), Roll No. 4408
I.A.S.R.I, Library Avenue, New Delhi – 110 012

Chairperson: Dr. Cini Varghese

Abstract: Lattice designs are resolvable incomplete block designs, some of which are Balanced Incomplete Block (BIB) designs or Partially Balanced Incomplete Block (PBIB) designs. These designs were developed for the comparison of large number of varieties in agricultural experiments. They can be broadly classified as square lattice designs, circular lattice designs, cubic lattice designs and rectangular lattice designs according to number of treatments, block size and number of restrictions imposed on randomization. Some methods of constructing these designs along with their association schemes are discussed here.

Key words: *Balanced Incomplete Block Design, Partially Balanced Incomplete Block Design, Association Scheme, Resolvable Design*

1. Introduction

When the number of treatments is large and/or complete blocks are unavailable or inappropriate, incomplete block designs are desirable and useful. Lattice design is nothing but the method of constructing certain types of resolvable incomplete block designs, some of which are Balanced Incomplete Block (BIB) designs or Partially Balanced Incomplete Block (PBIB) designs, but others are not. BIB designs require usually a large number of replications and are not available for all combinations of parametric values. Historically, lattice designs were developed for large-scale agricultural experiments (Yates, 1936) in which large number of varieties were to be compared. The main application since then has been and continues to be in agriculture.

A special feature of lattice designs is that the number of treatments, v, is related to the block size, k, in the form $v = k^2$ or $v = k^3$ or $v = k\,(k+1)$ or $v = k^2/2$. Even though this limits the number of possible designs, lattice designs represent an important class of designs. In certain types of agronomic or breeding experiments the number of treatments may be 100 or more.

2. Types of Lattice Designs

The existing lattice designs can be classified according to number of treatments, block size and number of restrictions imposed on randomization. They can be broadly classified as square lattice designs, circular lattice designs, cubic lattice designs and rectangular lattice designs. These designs are discussed in detail in Dey (1986), Varghese *et al*. (2004) and Hinkelmann and Kempthorne (2005).

2.1 Square Lattice Designs

The characteristic features of the square lattice designs introduced by Yates (1936) are that the number of treatments is a perfect square and the block size is the square root of this number. The number of replications of the treatments are flexible and these designs are

useful for situations in which a large number of treatments are to be tested. If the design has two replications of the treatments, it is called a simple lattice; if it has 3 replications it is called a triple lattice and so on. In general, if the number of replications is i, it is called an i-ple lattice design.

Square lattice designs can be constructed as follows:

Let there be $v = s^2$ treatments, numbered as 1, 2,…, s^2. Arrange these treatment numbers in the form of a $s \times s$ square array in natural order, i.e., in a standard array. The contents of each of the s rows of this array are taken to form blocks giving a set of s blocks; another set of s blocks forming another complete replication is obtained by taking the contents of each of the s columns of this array. We can check easily that this simple lattice design is a PBIB design with a Latin square association scheme, that is, a L_2-PBIB design. The parameters of this design are $v = s^2$, $b = 2s$, $r = 2$, $k = s$, $\lambda_1 = 1$, $\lambda_2 = 0$. Accordingly, two treatments are compared with variance V_1, say, if they occur together in the same block, V_2 otherwise.

Next, a $s \times s$ Latin square is taken and is superimposed on the above standard array of treatment numbers. The treatment numbers that fall on a particular symbol of the Latin square are taken to form a block. Thus s blocks corresponding to the s symbols of Latin square can be obtained. Again, another Latin square orthogonal to the previous one is taken and from this square also, another set of s blocks is obtained in the same manner. The process is repeated to get further replications. The process is continued till i-2 ($\leq$ s-1, if s is a prime or power of a prime) mutually orthogonal Latin squares are utilized. When the complete set of $(s-1)$ mutually orthogonal Latin squares (if such a set exists) is utilized, the design becomes a balanced $(s+1)$- lattice. A balanced lattice is a BIB design.

Thus the triple lattice represents an L_3 - PBIB(2) design with the parameters $v = s^2$, $b = 3s$, $r = 3$, $k = s$, $\lambda_1 = 1$, $\lambda_2 = 0$. The quadruple lattice is an L_4 - PBIB(2) design with parameters $v = s^2$, $b = 4s$, $r = 4$, $k = s$, $\lambda_1 = 1$, $\lambda_2 = 0$. In general the i-tuple lattice is an L_i - PBIB(2) design with parameters $v = s^2$, $b = is$, $r = i$, $k = s$, $\lambda_1 = 1$ and $\lambda_2 = 0$.

Example 2.1.1: Let $v = 3^2 = 9$, the standard array is as follows:

1 2 3
4 5 6
7 8 9

Considering all the four replications, a balanced lattice design which is a BIB design with $v = 9$, $b = 12$, $r = 4$, $k = 3$ and $\lambda = 1$ is obtained as given below:

Blocks											
Rep I			**Rep II**			**Rep III**			**Rep IV**		
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	4	7	1	2	3	1	2	3	1	2	3
2	5	8	4	5	6	6	4	5	5	6	4
3	6	9	7	8	9	8	9	7	9	7	8

If the last two replications are deleted, it becomes a simple lattice design. If last one replication is deleted, the design obtained is a triple lattice design which is a PBIB (2) design following L_3 association scheme.

2.2 Circular Lattice Designs

Circular lattice designs given by Rao (1956) are PBIB(3) designs with $v = 2n^2$ treatments arranged in n concentric circles and n diameters, giving rise to $2n^2$ lattice points on the circles. Each circle and each diameter has 2n points on them. These designs can be obtained by dualizing the group-divisible designs with 2n treatments in $2n^2$ blocks of size 2 each, the block contents of the GD designs being

$$(i, n+1), (i, n+2), ..., (i, 2n)$$

$$(i, n+1), (i, n+2), ..., (i, 2n), \quad i = 1, 2, ..., n$$

The parameters of design obtained are $v = 2n^2, b = 2n, r = 2, k = 2n, \lambda_1 = 2, \lambda_2 = 1, \lambda_3 = 0$.

Example 2.2.1: For n = 2, v = 8. The block contents of GD with v = 4, b = 8, r = 4, k = 2, $\lambda_1 = 0$, $\lambda_2 = 2$ are as given:

1	3
1	4
1	3
1	4
2	3
2	4
2	3
2	4

By taking dual of the above design, a circular lattice design with v = 8, b = 4, r = 2, k = 4, $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = 0$ is obtained and is given below:

1	2	3	4
5	6	7	8
1	3	5	7
2	4	6	8

This design can also be obtained by identifying the lattice points as treatments and circles and diameters as blocks.

Association Scheme

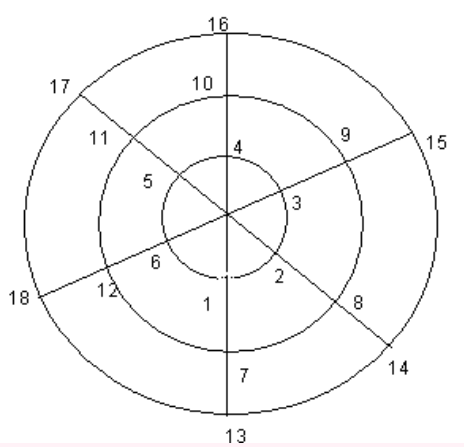
Corresponding to any treatment, the first associate is the treatment that is on the same circle and same diameter, second associates are those, which are either on the same circle or on the same diameter, and the rest are third associates {Rao, 1956}. Here,

$$v = 2n^2, n_1 = 4, n_2 = 4(n-1), n_3 = 2(n-1)^2,$$

$$\mathbf{P}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4(n-1) & 0 \\ 0 & 0 & 2(n-1)^2 \end{bmatrix}, \mathbf{P}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2(n-2) & 2(n-1) \\ 0 & 2(n-1) & 2(n-1)(n-2) \end{bmatrix} \text{ and}$$

$$\mathbf{P}_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 4 & 4(n-2) \\ 1 & 4(n-2) & 2(n-2)^2 \end{bmatrix}.$$

Example 2.2.2: Let $n = 3$ and $v = 2n^2 = 18$.



The various associates of treatment 1 are as follows:

1st Associates	2nd Associates	3rd Associates
4	7,13,16,10,2,3,5,6	8,9,11,12,14,15,17,18

2.3 Generalized Circular Lattice Designs

Circular lattice designs can be constructed easily, but the number of designs which can be obtained is very limited in this case. Hence generalized circular lattice designs were introduced (Varghese and Sharma, 2004) which covers more number of treatments.

Let the number of treatments be $v = 2sn^2$, $n \geq 2$. Draw n concentric circles and n diameters. Arrange the treatments on the points of intersection of these circles and the diameter such that s treatments occur at each point of intersection. Taking the circles and diagonals as blocks, we get a generalized circular lattice design which is a PBIB(3) design with parameters $v = 2sn^2$, $b = 2n$, $r = 2$, $k = 2sn$, $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = 0$.

Association Scheme

Two treatments are the first associates, if they lie on the same circle and on the same diameter, second associates, if they lie either on the same circle or on the same diameter, and third associates, otherwise. The parameters of the association scheme are seen to be:

$$v = 2sn^2, n_1 = 2s-1, n_2 = 4s(n-1), n_3 = 2s(n-1)^2, n \geq 2,$$

$$\mathbf{P_1} = \begin{bmatrix} 2s-1 & 0 & 0 \\ 0 & 4s(n-1) & 0 \\ 0 & 0 & 2s(n-1)^2 \end{bmatrix}, \mathbf{P}_2 = \begin{bmatrix} 0 & 2s-1 & 0 \\ 2s-1 & 2s(n-2) & 2s(n-1) \\ 0 & 2s(n-1) & 2s(n-1)(n-2) \end{bmatrix}$$

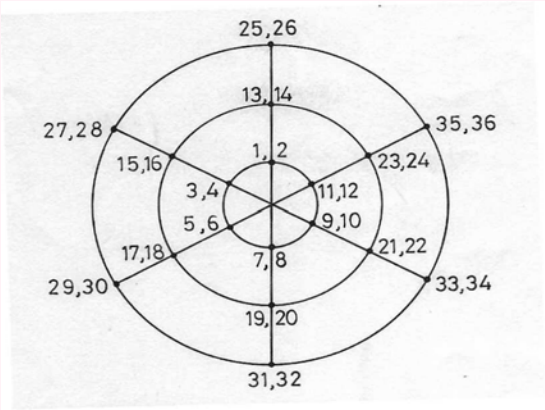
and

$$\mathbf{P_3} = \begin{bmatrix} 0 & 0 & 2s-1 \\ 0 & 4s & 4s(n-2) \\ 2s-1 & 4s(n-2) & 2s(n-2)^2 \end{bmatrix}.$$

Example 2.3.1: Let $n = 3$ and $s = 2$. The design with v = 36, b = 6, r = 2, k = 12, $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = 0$ is obtained as follows:

1	2	3	4	5	6	7	8	9	10	11	12
13	14	15	16	17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32	33	34	35	36
1	2	7	8	13	14	19	20	25	26	31	32
3	4	9	10	15	16	21	22	27	28	33	34
5	6	11	12	17	18	23	24	29	30	35	36

Example 2.3.2: Let s = 2, n = 3 and so v = 36. The arrangements of 36 treatments are as shown in the figure below:



The associates of treatment 1 are:

1st Associates	2nd Associates	3rd Associates
2,7,8	3,4,5,6,9,10,11,12,13,14,19,20, 25,26,31,32	15,16,17,18,21,22,23,24,27,28, 28,29,30,33,34,35,36

Note: When s = 1, this scheme reduces to circular lattice association scheme.

2.4 Cubic Lattice Designs

A cubic lattice design [Das and Giri, 1986] is a PBIB (3) design for $v = s^3$ treatments in blocks of size s and 3 replications, based on cubic association scheme. The s^3 treatments are represented by triplets $\alpha\ \beta\ \gamma$ ($\alpha, \beta, \gamma = 1, 2, \ldots, s$). Within each of the 3 replications, the

treatments are grouped into s^2 blocks, each of size s. In the first replication, the rule of this grouping is to keep the first two digits of the triplet constant within a block, allowing the last digit to take values from 1 to s. To form the blocks in the second replication, we keep the first and the last digits fixed within any block and giving the second digit all values from 1 to s. In the third replication, the second and last digits are kept constant within each block. This design has the parameters

$$v = s^3, b = 3s^2, \ r = 3, \ k = s, \ \lambda_1 = 1, \ \lambda_2 = 0 \text{ and } \lambda_3 = 0.$$

Example 2.4.1: Let $s = 3$ giving v = 27. To obtain a plan the 27 treatments are numbered by means of a three-digit code in which each digit takes all values from 1 to 3. For 27 treatments, the codes are as given in the following table:

Treatment No.	Code	Treatment No.	Code	Treatment No.	Code
1	111	10	211	19	311
2	112	11	212	20	312
3	113	12	213	21	313
4	121	13	221	22	321
5	122	14	222	23	322
6	123	15	223	24	323
7	131	16	231	25	331
8	132	17	232	26	332
9	133	18	233	27	333

The 9 group of treatments constitute the 9 blocks, block 1 containing the treatments (111),(112),(113). To form the blocks in the second replication, we keep the first and the last digits fixed within any block and given the second digit all values form 1 to 3. The first block therefore contains (111), (121), (131); the second block contains (112), (122), (132); and so on. In the third replication, the first and the second digits are constant within each block.

The design obtained is as follows with parameters v = 27 =b, r = 3 = k, $\lambda_1 = 1$, $\lambda_2 = 0$, $\lambda_3 = 0$:

Replication I			**Replication II**			**Replication III**		
1	2	3	1	4	7	1	10	19
4	5	6	2	5	8	2	11	20
7	8	9	3	6	9	3	12	21
10	11	12	10	13	16	4	13	22
13	14	15	11	14	17	5	14	23
16	17	18	12	15	18	6	15	24
19	20	21	19	22	25	7	16	25
22	23	24	20	23	26	8	17	26
25	26	27	21	24	27	9	18	27

Association Scheme

Let there be v= s^3 treatments denoted by (α,β,γ) $\alpha,\beta,\gamma = 1,2,...,s$. Define the distance δ between two treatments (α, β, γ) and (α', β', γ') to be the number of non-null elements in ($\alpha - \alpha'$, $\beta - \beta'$, $\gamma - \gamma'$). Two treatments are 1st, 2nd or 3rd associates according as δ = 1, 2, or 3 respectively {Raghavarao and Chandrasekhararao, 1964}. The parameters of this scheme are $n_1 = 3(s-1)$, $n_2 = 3(s-1)^2$, $n_3 = (s-1)^3$,

$$\mathbf{P}_1 = \begin{bmatrix} s-2 & 2(s-1) & 0 \\ 2(s-1) & 2(s-1)(s-2) & (s-1)^2 \\ 0 & (s-1)^2 & (s-1)^2(s-2) \end{bmatrix}, \mathbf{P}_2 = \begin{bmatrix} 2 & 2(s-2) & s-1 \\ 2(s-2) & 2(s-1)+(s-2)^2 & 2(s-1)(s-2) \\ s-1 & 2(s-1)(s-2) & (s-1)(s-2)^2 \end{bmatrix}$$

and

$$\mathbf{P}_3 = \begin{bmatrix} 0 & 3 & 3(s-2) \\ 3 & 6(s-2) & 3(s-2)^2 \\ 3(s-2) & 3(s-2)^2 & (s-2)^3 \end{bmatrix}.$$

Example 2.4.2: Let $s = 2, v = s^3 = 8$. The eight triplets are as given below:

Treatment	Triplets		
	α	β	γ
1	0	0	0
2	0	0	1
3	0	1	0
4	0	1	1
5	1	0	0
6	1	0	1
7	1	1	0
8	1	1	1

The three types of associates of the treatment 1 as given in the following table are obtained by taking all differences between (α,β,γ) and (α',β', γ').

1st Associates	2nd Associates	3rd Associates
2, 3, 5	4, 6, 7	8

2.5 Rectangular Lattice Designs

The class of two dimensional lattices is applicable when the number of treatments is a perfect square; i.e., v = s^2. Obviously, this limits the number of cases in which such a lattice design can be used. To remedy this deficiency Harshberger (1947, 1949) developed rectangular lattices for $s(s+1)$ treatments in blocks of size s units. These designs form a useful addition to the square lattices. The new designs are less symmetrical than the square lattices, in the sense that there is a greater variation in the accuracy with which two treatment means are compared. There are several methods of constructing rectangular lattice designs.

2.5.1 Simple Rectangular Lattice Designs

Denote the s(s+1) treatments by pair of symbols (x, y) such that $(x \neq y)$, $1 \leq x, y \leq (s+1)$. Blocks in the first replication are formed by selecting those treatments whose x-coordinate is same in the corresponding pairs of symbols and second replication is obtained by forming blocks with those treatments whose y-ordinate is same in the corresponding pairs of symbols. Nair (1951) has shown that these groups of treatments define an association scheme for a PBIB (4) design with. $v = s(s+1)$, $b = 2(s+1)$, $r = 2$, $k = s$.

Association Scheme

Two treatments (x, y) and (x', y') are said to be

1^{st} associates if $(x = x', y \neq y')$ or $(y = y', x \neq x')$,

2^{nd} associate if $(x \neq x', x \neq y', y \neq x', y \neq y')$,

3^{rd} associate if $(x = y', y \neq x')$ or $(y = x', x \neq y')$ and

4^{th} associate if $(x = y', y = x')$.

Thus the parameters of the scheme are: $n_1 = 2(s-1)$, $n_2 = (s-1)(s-2)$, $n_3 = 2(s-1)$, $n_4 = 1$, $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = \lambda_4 = 0$.

Furthermore,

$$P_1 = \begin{bmatrix} s-2 & s-2 & 1 & 0 \\ s-2 & (s-2)(s-3) & s-2 & 0 \\ 1 & s-2 & s-2 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 2 & 2(s-3) & 2 & 0 \\ 2(s-3) & (s-3)(s-4) & 2(s-3) & 1 \\ 2 & 2(s-3) & 2 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$P_3 = \begin{bmatrix} 1 & (s-2) & s-2 & 1 \\ (s-2) & (s-2)(s-3) & (s-2) & 0 \\ s-2 & (s-2) & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad P_4 = \begin{bmatrix} 0 & 0 & 2(s-1) & 0 \\ 0 & (s-1)(s-2) & 0 & 0 \\ 2(s-1) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

For $s = 2$ this design reduces to a PBIB (3) design since then $n_2 = 0$.

Example 2.5.1.1: $v = s(s+1) = 6$ treatments can be represented by the following ordered pairs:

Treatment Number	x	y
1	1	2
2	1	3
3	2	1
4	2	3
5	3	1
6	3	2

By selecting those treatments whose x-coordinate is same in the corresponding pairs of symbols and second replication is obtained by forming blocks with those treatments whose y-ordinate is same in the corresponding pairs of symbols. The design obtained is as follows with v = 6, b = 6, r = 2, k = 2:

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

2.5.2 Triple Rectangular Lattice Designs

Denote the treatments by triplets (x, y, z) with $x, y, z = 1, 2\ldots.s$ and $(x \neq y \neq z \neq x)$. The triplets are chosen in the following way.

Take a Latin square of order s+1 replacing the Latin letters A, B, C… by the Latin numbers 1, 2, 3,… respectively, arranged in such a way that the diagonal contains the numbers 1, 2,…, (s+1). The important feature of this arrangement is that no two treatments are in the same block more than once. The use of a Latin square with different letters down the leading diagonal ensures that in the third replication the three numbers associated with any letter are in the different rows and columns, and hence have not previously occurred together in a block. Then leaving out the diagonal, the resulting s(s+1) cells are identified by the row number x, the column number *y*, and the Latin number *z*. Each such cell corresponds to a treatment (x, y, z), and the treatments, which are allocated to the blocks in the three replicates as follows:

Replication I: Treatments with the same *x* value form the x^{th} block
Replication II: Treatments with the same *y* value form the y^{th} block
Replication III: Treatments with the same *z* value form the z^{th} block

This gives rise to a resolvable design with parameters v = s(s+1), b = 3s, r = 3, k = s. This method can be generalized to construct rectangular lattices with more than 3 replicates by using several Mutually Orthogonal Latin Squares (wherever available) to label the treatments appropriately.

Nair (1951) has shown that for $s = 2$ and $s = 3$ the resulting design is a PBIB design, but that this is no longer true for $s \geq 4$. For $s = 2$ a PBIB (2) design with the following association scheme exist: two treatments (x, y, z) and (x', y', z') are said to be 1^{st} associates if $x = x'$ or $y = y'$ or $z = z'$ and 2^{nd} associates, otherwise. It then follows that $n_1 = 3,\ n_2 = 2,\ \lambda_1 = 1,\ \lambda_2 = 0$ and

$$\mathbf{P}_1 = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \text{ and } \mathbf{P}_2 = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

For s = 3, the association is as follows: two treatment (x, y, z) and (x', y', z') are said to be first associate if $x = x'$ or $y = y'$ or $z = z'$ and 2^{nd} associates, if they have three digits

alike, and 3rd associates, otherwise. This leads to a PBIB (3) design with parameters $n_1 = 6,\ n_2 = 2,\ n_3 = 3,\ \lambda_1 = 1,\ \lambda_2 = \lambda_3 = 0$ and

$$\mathbf{P}_1 = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix},\ \mathbf{P}_2 = \begin{bmatrix} 3 & 0 & 3 \\ 0 & 1 & 0 \\ 3 & 0 & 0 \end{bmatrix} \text{ and } \mathbf{P}_3 = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Example 2.5.2.1: Form a 4 × 4 square as follows:

A	B1	C2	D3
D4	C	B5	A6
B7	A8	D	C9
C10	D11	A12	B

In the first replication, place in a block all numbers that lie in the same row of the Latin square, in the second replication all numbers that lie in the same column, and in the third all numbers that have the same Latin letter.

The resultant design is:

1	2	3
4	5	6
7	8	9
10	11	12
4	7	10
1	8	11
2	5	12
3	6	9
6	8	12
1	5	7
2	9	10
3	4	11

By using the first 2 replications from any plan we obtain a rectangular lattice in 2 replications is obtain, which by means of replications can be used for an experiment in 4, 6, 8, etc., replications. By using all 3 replications of the plan we have the designs for 3, 6, 9, etc., replications.

2.5.3 Method by Sharma and Das (1985)

Let v = ks be the number of treatments denoted by 1, 2,…,ks. These treatments are arranged in a rectangle R with k rows and s columns. A set of r MOLR of order k × s in s symbols is formed. A Latin rectangle is now taken from the set and is superimposed on the rectangle R. The k treatment numbers that fall on a symbol of the Latin rectangle are taken to form a block. Thus from the s symbols s blocks are obtained which constitutes the first

replication. More replications are obtained using more MOLR. When s is prime or prime power, lattice designs obtained is PBIB(3) following rectangular association scheme.

Example 2.5.3.1: Let v = ks =15. These treatments can be written in the following rectangle as:

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15

A complete set of MOLR in 3 rows and 5 columns each is as given below:

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

The rectangular lattice design corresponding to first two MOLR is as follows:

Rep I			Rep II		
1	10	14	1	9	12
2	6	15	2	10	13
3	7	11	3	6	14
4	8	12	4	7	15
5	9	13	5	8	11

Parameters of the above design are v = 15, b = 10, r = 2 and k = 3.

2.5.4 Another Method

As pointed out by Shrikhande (1965), rectangular lattice designs can also be constructed from a balanced lattice with $(s+1)^2$ treatments. If the first replication is omitted, and if all treatments that appear in any one selected block in the first replicate are omitted the resultant design is a rectangular lattice in which every block is of size s and in which no two treatments appear together more than once in the same block. This method gives any number of replicates up to s, though of course the method fails when no balanced lattice exists, as with 36 treatments.

Example 2.5.4.1: Consider a balanced lattice s with $v = (s+1)^2 = 3^2$.

Blocks											
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	*4*	~~*7*~~	1	2	3	1	2	3	1	2	3
2	*5*	~~*8*~~	4	5	6	6	4	5	5	6	4
3	*6*	~~*9*~~	~~7~~	~~8~~	~~9~~	~~8~~	~~9~~	~~7~~	~~9~~	~~7~~	~~8~~

The rectangular lattice design obtained is:

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

Parameters are $v = s(s+1) = 6, b = (s+1)^2 = 9, r = (s+1) = 3, k = s = 3$.

3. Conclusions

Lattice designs form an important class of useful resolvable incomplete block designs. These designs were initially developed for the comparison of large number of varieties in agricultural experiments. They can be broadly classified as square lattice designs, circular lattice designs, cubic lattice designs and rectangular lattice designs. In square lattice designs, the number of treatments is a perfect square and the block size is the square root of this number. The number of replications of the treatments is flexible in these designs and is useful for situations in which a large number of treatments are to be tested. Some of these designs are BIB designs while others are PBIB(2) designs. Circular lattice designs are PBIB(3) designs with the treatment structure as $v = 2n^2$. These designs can be constructed easily, but the number of designs which can be obtained is very limited in this case. To overcome this, generalized circular lattice designs were obtained which are again PBIB(3) designs. Cubic lattice designs were introduced for plant breeding experiments in which selection are to be made from an unusually large number of varieties. The number of treatments in this case must be an exact cube and the size of the block is the cube root of the number of treatments. Thus cubic lattices can accommodate a large number of treatments in small size of incomplete block. The number of replications must be 3 or a multiple of 3. The treatment structure of rectangular lattice designs is of the form s(s+1) and the number of replications is flexible. These designs are generally higher associate class PBIB designs and hence there is a greater variation in the precision with which two treatments are compared.

References

Das, M.N. and Giri, N.C. (1986). *Design and Analysis of Experiments*. Wiley Eastern Limited, New Delhi.

Dey, A. (1986). *Theory of Block Designs*. Wiley Eastern Ltd., New Delhi.

Harshbarger, B.(1947). Rectangular lattices. *Va. Agri. Expt. Stat. Memoir*, **1**,1-26.

Harshbarger, B.(1949). Triple rectangular lattices. *Biometrics*, **5**, 1-23.

Hinkelmann, K. and Kempthorne, O. (2005). Design and Analysis of Experiments. Wiley Interscience Publications, New Jersey.

Nair, K.R. (1951). Rectangular lattices and partially balanced incomplete block designs. *Biometrics*, **7**, 145-154.

Raghavrao, D. and Chandrasekhararao, K. (1964). Cubic designs. *Ann. Math. Statist*., **35**, 389-397.

Rao, C.R. (1956). A general class of quasifactorial and related designs. *Sankhya*, **17**, 165-174.

Sharma, V.K. and Das, M.N. (1985). On resolvable incomplete block designs. *Austral. J. Statist*., **27**(3), 298-302.

Shrikhande, S.S. (1965). On a class of partially balanced incomplete block designs. *Math. Statist*., **36**, 1807-1814.

Varghese, Cini and Sharma, V.K. (2004). A series of resolvable PBIB(3) designs with two replicates. *Metrika*, **60**, 251-254.

Varghese, Cini, Sharma, V.K., Jaggi, Seema and Sharma, Anu. (2004). Three-associate class partially balanced incomplete block designs and their application to partial diallel crosss. Project Report, IASRI Publication, New Delhi.

Yates, F. (1936b). A new method of arranging variety trial involving a large number of varieties. *J. Agri. Sci*., **26**, 424-455.

BIOASSAYS

R.Srivastava
I.A.S.R.I., Library Avenue, New Delhi-110012
ravindra@iasri.res.in

1. Introduction

Biological assays (*bioassays*) are a set of techniques relevant to the comparisons between the strength of alternative but similar biological stimuli (a pesticide, a fungicide, a drug, a vitamin, plant extract, etc.) based on the response produced by them on the subjects (e.g., an animal, a piece of animal tissue, a plant, a bacterial culture, subhuman primates or humans, living tissues, plants or isolated organisms, insects, etc). Typically, a bioassay involves (i) *stimulus,* (ii) *subject,* and (iii) *response,* the change produced on the subject due to application of stimulus (such as an analytical value like blood sugar content or bone ash percentage, occurrence or non-occurrence of a certain muscular contraction, recovery from symptoms of a dietary deficiency, or death, etc).

Normally, two preparations of the stimulus, one of known strength (*standard preparation*) and another of unknown strength (*test preparation*), both with quantitative doses, are applied to a set of living organisms. The general objective of the bioassays is to draw statistically valid conclusions on the *relative potency* of *test preparation* with respect to *standard* one. If d_s and d_t denote the doses of the standard and the test preparations respectively such that each of them produces a pre-assigned response in some living organism, then the ratio $\rho = d_s / d_t$ is called the *relative potency* of the test preparation. If ρ is greater than unity, it shows that a smaller dose of the *test preparation* produces as much response as relatively larger dose of *standard preparation* and therefore the potency of the *test preparation* is greater than that of *standard preparation.* Similarly, when ρ is less than unity, the potency of the *standard preparation* is greater than that of the *test preparation*. Naturally, such statistical procedures may depend on the nature of the stimulus and response as well as on other extraneous experimental (biological or therapeutical) considerations.

As is clear from the above discussion, the biological assay is an experiment in which the interest lies in comparing the potencies of the treatments on an agreed scale. Biological assays are. therefore, different from traditional comparative experiments where the interest lies in comparing the magnitude of effects of treatments. The experimental technique may be same, but the difference in purpose affects the designing and the statistical analysis of the experimental data. Thus, an investigation into the effects of different samples of insulin on blood sugar of rabbits is not necessarily a biological assay; it becomes one if the interest lies not simply in the change in blood sugar levels, but in their use for the estimation of the samples on a scale of standard units of insulin. A field trial of the responses of the potatoes to various phosphatic fertilizers would not generally be regarded as an assay; but if the yields of potatoes are to be used in assessing the potency of a natural rock phosphate relative to a standard superphosphate, and perhaps even in estimating the availability of phosphorus in the rock phosphate, then the experiment is an assay.

A glance at the history and day to day life reveals that the bioassays were (are) commonly employed, though sometimes unknowingly. We, therefore, give a brief account of bioassays with historical perspective.

1.1 History and Some Common Practices

The earliest description of biological assay can be found in the Bible wherein Noah sends a dove from his ark till it returns with an olive leaf. By doing this experiment, Noah estimates the levels of receding waters from the Earth's ground.

This experiment has all the three essential components of an assay- namely, 'stimulus' (depth of water), 'subject' (the dove) and 'response' (plucking of olive leaf). Knowledge of the response enabled Noah to estimate, or rather, in this instance, to place an upper limit to the size of the stimulus. The limitations of his animal house made his replication less than would today be thought as adequate, but in other respects his experiment was admirable for its purpose.

Figure1.1: Canary
Source:http://www.photos-animoux.com

In bioassays bio-organism is used to test toxicity of chemicals. One can notice that the bioassays are employed (may be unknowingly) in real life. Traditionally, coal miners have taken caged canaries down into the mines to help ensure a safe air supply. Canaries are more sensitive than humans to methane, so they were used to provide an advanced warning of rising levels of methane to dangerous levels in the mines. Death of canary meant warning signal to miners and they left the mine immediately.

Serious work on biological assays began by the end of 19^{th} century with Ehrlich's investigations into the standardization of diphtheria antitoxin. Since then it became a common practice to standardize materials by means of the reactions on living organism. This has found favour in many branches of science like pharmacology and plant pathology. The assays were, however, put on sound footing only since 1930's when some statisticians contributed through their refined methods.

1.2 Types of Bioassays

The *test preparation* behaves (in terms of the response/tolerance distribution) as if it is a dilution or concentration of the *standard preparation*. These assays are termed as *dilution assays* or *analytical dilution assays.* Analytical dilution assays can broadly be classified into

two categories viz., (i) direct dilution and (ii) indirect dilution assays. There is one more category of bioassays viz. qualitative assay which do not pose any statistical problems.

1.2.1 Direct Assays are those assays where dose needed to produce a pre-assigned/ specified response is directly measurable for both the preparations. In this case the response is certain while the dose is a non-negative random variable that defines the tolerance distribution. These assays are practical only when it is possible to administer the dose in such a manner that the minimal amount of dose to produce a specified response can be measured directly. Following example given by Finney (1978*, Statistical Methods in Biological Assays, 3rd Edition*) makes the ideas clear.

Example 1: (Burn, Finney and Goodwin 1950; Hatcher and Brody 1910). This is a typical example of direct assay '*the cat method'* for the assay of digitalis. The standard or test preparation is infused at a fixed rate, into the blood stream of a cat until the heart stops beating. The total time of infusion multiplied by rate is termed as dose. This is repeated on several cats for each preparation and the mean doses are compared *i.e.*, if $\bar{x}_s$ and $\bar{x}_t$ denote average of tolerances for *standard* and *test* preparations respectively then an estimate of relative potency is given by $R = \bar{x}_s / \bar{x}_t$. Three groups of cats were infused with two tinctures of Strophantus using the procedure described above. The tinctures were having same effective ingredients. The doses were recorded as quantities per kg body weight of cat.

Table 1.1 shows the fatal doses or *tolerances* of three groups of cats for two tinctures of strophanthus (Burn *et al*., 1950). The doses were recorded as quantities per kg body weight of cat, and unfortunately the total doses are not available: the tolerance is thus assumed to vary in proportion to body weight, or at least to show an approximately proportion variation rather than independence of body weight. Provided that the cats have been assigned at random to the different preparations, either form of expression of dose gives a valid method of estimating potency, but neither necessarily makes the best possible use of information on body weight.

Table 1.1: Tolerances of cats for tinctures of strophantus and ouabain

Preparation	Strophanthus A (in 0.01 c.c per kg.)	Strophanthus B (in 0.01 c.c per kg.)
Tolerances ...	1.55	2.42
	1.58	1.85
	1.71	2.00
	1.44	2.27
	1.24	1.70
	1.89	1.47
	2.34	2.20
	--	--
	--	--
Mean …	1.68	1.99

Suppose that Strophanthus tincture B is to be regarded as the standard preparation, and A is to be compared with it as a test preparation. From the means in Table 1.1, 0.0168 c.c. of A is

estimated to produce the same results as 0.0199 c.c. of B, either being just sufficient, on an average, to kill a cat. Hence the relative potency is estimated to be

$$R = \hat{\rho} = \frac{0.0199}{0.0168} = 1.18;$$

Having obtained the estimate of relative potency ρ as above, a natural question is that of its precision. We now work out the precision of estimate.

Let $R = \bar{x}_B / \bar{x}_A$, then variance of R is given by

$$\mathrm{V}(R) = \frac{1}{\bar{x}_A^2}\left[\mathrm{V}(\bar{x}_B) + R^2\,\mathrm{V}(\bar{x}_A)\right]$$

Now $\sum (x_i - \bar{x})^2 = 0.7587$ for A and 0.6815 for B. It can be seen that both the sums of squares are based on 6 degrees of freedom and, therefore, estimate of common variance is given by

$s^2 = \dfrac{0.7587 + 0.6815}{12} = 0.1200$ and

$\mathrm{V}(\bar{x}_A) = \mathrm{V}(\bar{x}_B) = \dfrac{s^2}{7}$ (since both $\bar{x}_A, \bar{x}_B$ are based on 7 observations) Here, $R = 1.18$ and

therefore $\mathrm{V}(R) = \dfrac{s^2}{\bar{x}_A^2}\left[\dfrac{1}{7} + \dfrac{R^2}{7}\right] = 0.0145$.

Hence, $R = 1.18 \pm 0.120$ since $\sqrt{0.0145} = 0.120$.

We now attempt to work out the fiducial limits of the estimate of relative efficiency R using Fieller's Theorem that is given below

Theorem 1.1: Let α and β be two parameters with $\mu = \alpha/\beta$ such that $E(a) = \alpha$ and $E(b) = \beta$. Let $m = a/b$ be an estimate of μ. We also assume that a and b are normally distributed. Also, $\mathrm{V}(a) = v_{11}s^2, \mathrm{V}(b) = v_{22}s^2, \mathrm{Cov}(a,b) = v_{12}s^2$, where s^2 is an unbiased estimator for common variance σ^2 based on f degrees of freedom. Under this set up upper and lower fiducial limits of μ are given by

$$m_L, m_U = \frac{1}{(1-g)}\left[\left(m - \frac{g v_{12}}{v_{22}}\right) \pm \frac{t_0 s}{b}\left\{v_{11} - 2m v_{12} + m^2 v_{22} - g\left(v_{11} - \frac{v_{12}^2}{v_{22}}\right)\right\}^{1/2}\right]$$

where $g = t_0^2 s^2 v_{22} / b^2$, and t_0 is the two-sided α percent point of Student's t distribution with f degrees of freedom.

We can very easily use Fieller's Theorem to calculate fiducial limits of R of Example 1. We have

$s^2 v_{11} = s^2 v_{22} = \frac{0.1200}{7}, v_{12} = 0$ and hence $g = \frac{(2.18)^2 \times 0.120}{(1..68)^2 \times 7} = 0.0289$ and, therefore, the fiducial limits are

$$R_L, R_U = \left[1.18 \pm \frac{2.18}{1.68} \left\{ \frac{0.1200}{7} \times \left(0.9711 + 1.18^2 \right)^{1/2} \right\} \Big/ 0.9711 \right] = 0.95, 1.48$$

An analysis in terms of log doses may be more satisfactory than one in the absolute units. In direct assays, the assumption of a normal distribution of log tolerances, if admissible, has many advantages. All variance estimates may be pooled in order to give the best possible estimate of the population variance. The relative potency is estimated as the antilogarithm of the difference of two means, instead of as a ratio of two means, fiducial limits are calculated from simple standard error formulae without the use of Fieller's Theorem.

Direct assays many a times pose serious problems that may lead to erroneous conclusions. For example if two stimuli of unequal potency are applied at equal rates (e.g. two drugs infused with equal speeds), subjects receiving the less potent stimulus will have longer average times under treatment than those receiving the more potent stimulus. If there is any time lag in the production of effects on the subject, or any cumulative effect other than of a simply additive nature, the comparison of the two drugs will be biased. Even when this danger is absent, technical difficulties may prevent the experimenter from ensuring that subjects receive just the right dose to produce the specified response; to determine individual tolerances of cats for digitalis may require no more than reasonable skills and care, but to determine individual tolerance of aphids for an insecticide is impossible, and a different method of assay must be sought. This leads us to indirect assays.

1.2.2 Indirect Assays: In most of the bioassays, response is not directly measurable and therefore indirect methods are used to estimate the dose corresponding to a given response via a *dose-response* relationship. These kinds of assays are known as *indirect assays.* In these assays the dose is administered at some prefixed (usually non-stochastic) levels, and at each level the response is observed for subjects included in the study. Thus, the dose is generally non-stochastic and the responses are stochastic in nature. The stochastic response provides information about the tolerance distribution of a particular preparation. If the response is a quantitative variable (magnitude of some property like survival time, weight, etc.), then we have *quantitative* assay. On the other hand if the response is *quantal* (i.e., all or nothing), we have *quantal* assay. Both these assays are commonly adopted in statistical practice. Within this framework, the nature of dose-response regression may call for suitable transformation on the dose variable (called the *dosage* or *dose-metameter*) and/or response variable called the *response metameter.* The objective of these transformations is to achieve a linear dose-response regression that may induce simplification in statistical modeling and analytical techniques. If z represents the dose in the original scale, then the two transformations that have been found useful in bioassays are (i) $x = log_e(z)$ and (ii) $x = z^{\lambda}$, where λ>0 is a known constant. The first of these gives rise to *parallel line assays* and the second to *slope ratio assays.* These assays generally fall in the category of quantitative indirect assays. Transformation of response variable is generally not needed in such bioassays. In quantal

assays, the response variable is generally subjected to the *probit* (or normit) and logit transformation based on normal and logistic distributions respectively.

In the sequel we describe, the indirect assays with quantitative response. Consider two assays A and B, each administered at $m(\geq 2)$ prefixed levels (doses) $d_1, d_2, ..., d_m$. Let $Y_{s_i}\left(Y_{t_i}\right)$ be the response variable of standard (test) preparations. It is not necessary to have the same number of doses for both the preparation, but the modifications are straight forward and hence we assume this congruence. We first assume that both Y_{s_i} and Y_{t_i} are continuous (and possibly non-negative) random variables. Suppose that there exist some dosage $x_i = \xi_i(d_i), i = 1,2,...,m$ and response metameter $Y^* = g(X)$ for some strictly monotone $g(.)$, such that the two dosage-response regressions may be taken as linear, namely

$$\begin{aligned} Y_{t_i}^* &= \alpha_t + \beta_t x_{t_i} + e_{t_i} \\ Y_{s_i}^* &= \alpha_s + \beta_s x_{s_i} + e_{s_i} \end{aligned} \tag{1.3.1}$$

where, for statistical inferential purposes, certain distributional assumptions are needed for the error components e_{t_i} and e_{s_i}, $i = 1,2,\ldots,m$. Generally, in the context of log dose transformations we have a parallel-line assay, while slope-ratio assays arise typically for *power transformations*.

1.2.2(a) Parallel Line Assays: For linearizing transformation, $x_i = \log(dose)$, $i = 1,2,\ldots,m$, let

$$E(Y_{s_i}^*) = \alpha_s + \beta x_{s_i} \tag{1.3.2}$$

denote the relation between the expected response and x_s where $x_s = \log(d_s)$ and d_s denotes the dose of the standard preparation. Denoting by d_t a dose equipotent to d_s, we have $\rho = d_s / d_t$, that is

$$\log \rho = \log d_s - \log d_t = x_s - x_t .$$

That is $x_s = \log \rho + x_t$. Substituting for x_s in the relation of the standard preparation (1.3.2), we get the relation for the test preparation as

$$E\left(Y_{s_i}^*\right) = \alpha_s + \beta\left(\log \rho + x_{t_i}\right)$$

$$\text{that is} \quad E\left(Y_{s_i}^*\right) = \alpha_t + \beta x_{t_i} = E\left(Y_{t_i}^*\right) \tag{1.3.3}$$

where $\alpha_t = \alpha_s + \beta \log \rho$. Hence, the relationship for the test preparation is also linear like that of the standard preparation for the same transformation. An examination of the two equations for the two preparations shows that the lines have the same slope and are, therefore, parallel.
In this setup we then have $\beta_s = \beta_t = \beta$ (unknown), while, $\alpha_t = \alpha_s + \beta \log \rho$, where ρ is the relative potency of the test preparation with respect to the standard one. This leads to the basic estimating function

$$\log \rho = (\alpha_t - \alpha_s)/\beta \tag{1.3.4}$$

so that if the natural parameters β, α_s and α_t are estimated from the acquired bioassay data set, statistical inferences on log ρ (and hence ρ) can be drawn in a standard fashion. If in an assay *m* doses are taken for each of the two preparations and $\bar{x}_s$ and $\bar{x}_t$ denote the averages of the dose metameters and $\bar{y}_s$ and $\bar{y}_t$ are the average responses for the preparations, then it is known that

$$\alpha_s = \bar{y}_s - \beta \bar{x}_s$$

and

$$\alpha_t = \bar{y}_t - \beta \bar{x}_t . \tag{1.3.5}$$

Substituting these values in log $\rho = (\alpha_t - \alpha_s)/\beta$, we get an estimate *R* of ρ from

$$\log R = \bar{x}_s - \bar{x}_t - \{\bar{y}_s - \bar{y}_t\}/\beta . \tag{1.3.6}$$

From equations (1.3.2) and (1.3.3) it is seen that the two lines for the two preparations should be parallel when the dose metameter is log (dose). The assays corresponding to this transformation are, therefore, called parallel line assays. Thus, in a parallel-line assay, the two dose-response regression lines (1.3.1) are taken to be parallel and, further that the errors e_{t_i} and e_{s_i} have the same distribution (often taken as normal). For normally distributed errors, the whole set of observations pertains to a conventional linear model with a constraint on the two slopes, β_s and β_t, so that the classical maximum likelihood estimators and allied *likelihood ratio tests* can be incorporated for drawing statistical conclusion on the relative potency or the fundamental assumption of parallelism of the two regression lines. The estimator of log ρ involves the ratio of two normally distributed statistics, and, therefore, it may be biased; moreover, generally the classical *Fieller's theorem* (see Finney, 1964) is incorporated for constructing a confidence interval for log ρ (and hence, ρ). Because of this difference in setups (with that of the classical linear model), design aspects for such parallel-line assays need a more careful appraisal. For equi-spaced (log) doses, a symmetric *2m*-point design has optimal information contents, and is more popularly used in practice. We refer to Finney (1964) for a detailed study of such bioassay design in a conventional normally distributed error model. Two main sources of non-robustness of such conventional inference procedure are the following:

1. Possible non-linearity of the two regression lines (they may be parallel but yet curvilinear)
2. Possible non-normality of the error distributions.

On either count the classical theory procedures may perform quite non-robustly and their (asymptotic) optimality properties may not hold even for minor departures from either postulation. If the two dose-response regressions (linear or not) are not parallel, then the fundamental assumption of parallel-line assays is vitiated, and hence statistical conclusion based on the assumed model may not be very precise

1.2.2(b) Slope Ratio Assays: For linearizing transformation, $x_i = (\text{dose})^{\lambda}$ $i = 1, 2, \ldots, m$, let

$$E\left(Y_{s_i}^{*}\right) = \alpha_s + \beta_s x_{s_i} \tag{1.3.7}$$

denote the relation between the expected response and x_s, where $x_s = (d_s)^\lambda$ and d_s denotes the dose of the standard preparation..

Denoting by d_t a dose equipotent to d_s, we have $\rho = d_s / d_t$, or $\rho^\lambda = (d_s/d_t)^\lambda$ that is

$$\rho^\lambda = \frac{x_s}{x_t} \tag{1.3.8}$$

That is, $x_s = x_t \rho^\lambda$. Substituting for x_s in the relation of the standard preparation (1.3.2), we get the relation for the test preparation as

$$E\left(Y_{s_i}^*\right) = \alpha_s + \beta_s x_t \rho^\lambda$$

$$\text{or, } E\left(Y_{s_i}^*\right) = \alpha_s + \beta_t x_{t_i} = E\left(Y_{t_i}^*\right) \tag{1.3.9}$$

where $\beta_t = \beta_s \rho^\lambda$, i.e., $\rho^\lambda = \beta_t / \beta_s$.

$$\rho = \{\beta_t / \beta_s\}^{1/\lambda} \tag{1.3.10}$$

Since the relative potency is estimated from the ratio of the slopes of the two preparations, the assays, corresponding to the transformation z^λ are called slope ratio assays.

The relative potency is typically a non-linear function of the two shapes β_t and β_s, and presumes knowledge of λ. In such a case the two error components may not have the same distribution even if they are normal. This results in heteroscedastic linear model (unless ρ =1), where the conventional linear estimators or allied tests may no longer possess validity and efficiency properties.

Since ρ^λ is a ratio of two slopes, its conventional estimator based on the usual estimators of the two slopes is of the ratio type. For such ratio-type estimators, again the well-known Fieller theorem is usually adopted to attach a confidence set to ρ or to test a suitable null hypothesis. Such statistical procedures may not have the exact properties for small to moderate sample sizes. Even for large sample sizes, they are usually highly non robust for departures from the model based assumptions (*i.e.* linearity of regressions, the fundamental assumptions of normality of errors.). Again, the design aspects for such slope ratio assays need careful study and Finney (1964) contains a detailed account of this study. Because of the common intercept, usually a $2m+1$ point design is advocated here.

1.2.2 (c) Quantal Assays: In quantal assays, occurrence or non-occurrence depends on the intensity of the stimulus. Under the controlled conditions, for any one subject there will be a certain level of intensity below which the response does not occur and above which the response occurs. This value is known as *threshold* or *limen,* but the term *tolerance* is widely accepted now. This tolerance value will vary from one member to another of the population used, frequently between quite wide limits. When the characteristic response is quantitative, the stimulus intensity needed to produce a response of any given magnitude will show similar variation between individuals. In either case, the value for an individual also is likely to vary from one occasion to another as a result of uncontrolled internal or external conditions.

In these assays, the earlier attempts were made to characterize the effectiveness of a stimulus in relation to a quantal response referred to as the *minimal effective dose* or for a more restricted class of stimuli, as the *minimal lethal dose* terms which failed to take account of the variation in tolerance within a population. The logical weakness of such concepts is the assumption that there is a dose for any given chemical which is only just sufficient to kill all or most of the insects of a given species, and, doses a bit lesser would not kill any insect of that species. Any worker, however, accustomed to the estimation of toxicity knows that these assumptions do not represent the reality.

It might be thought that the *minimal lethal dose* of a poison could instead be defined as the dose just sufficient to kill a member of the species with the least possible tolerance, and also a *maximal non-lethal dose* as the dose which will just fail to kill the most resistant member. Undoubtedly some doses are so low that no test subject will succumb to them and others so high as to prove fatal at all, but considerable difficulties attend determination of the end-points of these ranges. Even when the tolerance of an individual can be measured directly, to say from measurements on a sample of ten or a hundred that the lowest tolerance found indicated the *minimal lethal dose* would be unwise: a larger sample might contain a more extreme member. When only quantal responses for selected doses can be recorded the difficulty is increased, and the occurrence of exceptional individuals in the batches at different dose levels may seriously bias the final estimates. The problem is in fact that of determining the dose at which the dose response curve for the whole population needs the 0% or 100% levels of kill and even a very large experiment could scarcely estimate these points with desired accuracy.

An escape from the dilemma can be made by giving attention to a different and more satisfactorily defined characteristics, the *median lethal dose*, or, as a more general term to include response other then death, the *median effective dose*. This is the dose that will produce a response in half the population. The median effective dose is commonly referred to as the ED 50, the more restricted concept of median lethal dose as the LD 50. Analogous symbols were used for doses effective for other proportions of the population, ED90 being the dose that causes 90% to respond. With a fixed total number of subjects, effective doses in the neighborhood of ED50 can usually be estimated more precisely than those for more extreme percentage levels and this is, therefore, particularly favoured in expressing the effectiveness of the stimulus. The ED50 alternatively may be regarded as the median of the tolerance distribution that is to say the level of tolerance such that exactly half the subjects lie on either side of it. The ED 50 or LD 50 can easily be calculated using the *Probit Analysis*.

1.2.2 (d) Probit Analysis: We consider an experiment conducted with different doses of an insecticide under standardized conditions to samples of an insect species. Data on number of insects killed and the number exposed to the insecticide are recorded. Ratio of number of insects killed to number exposed gives the empirical probability of the insects killed at particular dose. This empirical probability data is subjected to probit or logit transformation and the doses are subjected to logarithmic transformation.

Probit transformation is obtained by adding 5 to the normal equivalent deviate. In this transformation, we replace each of the observed proportions with the value of standard normal

curve below which the observed proportion of the area is found. To avoid negative numbers, the constant 5 is usually added. For example, if half (0.5) of the subjects respond at a particular dose, the corresponding probit value is 0, since half of the area in a standard normal falls below a Z score of 0. When the constant 5 is added, the transformed value for the proportion is 5. If the observed proportion is 0.95, the corresponding probit value is 1.64. Addition of the constant value of 5 makes this 6.64. Likewise, if 10% of the subjects respond, then the normal equivalent deviate is -1.29 and hence the probit value is 3.7.

In the logit transformation, the observed proportion P is changed to

$$\frac{\ln\left(P/(1-P)\right)}{2}+5$$

The quantity $\ln\left(P/(1-P)\right)$ is called a logit. Division by 2 and addition of the constant 5 is done to keep the values positive and to keep the two types of transformations on a similar scale. If the observed proportion is 0.5, the logit-transformed value is 0+5=5, the same as the probit-transformed value. Similarly, if the observed proportion is 0.95, the logit-transformed value is 6.47 (=1.47+5). This differs somewhat from the corresponding probit value of 6.64. (In most situations, analyses based on logits and probits give very similar results.)

The above is a discussion about the transformation of the observed proportions. The doses are transformed to the logarithmic scale. When the experimental data on the relation between dose and proportion of the subjects responded have been obtained, either a graphical or a statistical approach in terms of fitting of response metameter-dose metameter linear regression relationship can be used to estimate the parameters such as regression coefficients and their standard errors. We can also obtain observed and expected frequencies, Pearson's goodness of fit chi-square, and confidence intervals for effective levels of independent variable(s). If the Pearson's goodness of fit chi-square is non-significant, we compute the value of the response metameter (probit) is 5. The antilogarithm of this dose metameter is the ED50. The effectiveness of stimulus is not fully described by ED50 alone. Two insecticides/fungicides may require same rate of application to be lethal for half the population. 'Spread' of tolerance distributions is very important. A stimulus with less spread will be more effective than with the larger spread. It is, therefore, very much necessary to calculate standard error and fiducial limits associated with ED50 or LD50. Besides knowing the ED50 of particular chemical preparation, the experimenter may be interested in comparing the relative potencies of several preparations with a standard preparation. For this, one is required to fit the probit regression lines and test for parallelism. If the probit regression lines are parallel than the relative potency will be constant at all levels of response.

Remark: The probit analysis can also be very useful in market research. Suppose, we want to study "How low does a sale price have to be fixed to induce a consumer to buy a product? Here the stimulus is the sale price of an object and the response is sale made or not. Since all the people who are going for shopping do not respond in the same way *i.e.* they have different tolerance levels. Therefore, one has to obtain the proportion of the shoppers responding to each level of price. Therefore, here the subjects are the shoppers.

1.3 Block Designs and bioassays

The contrasts to be estimated in bioassays are different from those in usual experiments where the interest is mainly in making all possible pairwise treatment comparisons or test treatment control treatment comparisons. In these experimental situations, the interest is only in limited number of treatment contrasts. For example in parallel line assays the contrasts of major interest are *preparation, combined regression* and *parallelism* contrasts while for slope ratio assays the contrasts of major interest are *blank* and *intersection contrasts*. For experiments involving bioassays, therefore, the usual classical designs are not appropriate as it is generally desired that the contrasts of interest be estimated free from block effects. If the number of homogeneous experimental units are same as the number of doses then the experiment is conducted in randomized complete block design. However, when the number of experimental units within homogeneous set are less than that of total number of doses then recourse is made to use of incomplete block designs. Thus, for the conduct of bioassays one should choose a block design that estimates contrasts of interest free from block effects.

2.1 Analytical Techniques for Bioassays

As mentioned above the block designs which are efficient for estimating all the elementary contrasts, are generally not suitable for the experiments involving bioassays, as the treatment contrasts of interest are different. In parallel line assays the contrasts to be estimated are preparation, combined regression and parallelism contrasts while for slope ratio assays the contrasts of interest are blank and intersection contrasts. In bioassays we seek to obtain efficient block designs for two types of bioassays, i.e., (i) parallel line assays (ii) slope ratio assays. We noe give a brief outline of analytical techniques relevant for biological assays. We begin with parallel line assays.

2.1.1Definition of contrasts for parallel line assays

Let s and t denote typical doses of standard and test preparations respectively with their effects being denoted by $\eta_1(x)$ and $\eta_2(z)$ respectively, where $x = log_e s,\ z = log_e t$. Let there be m_1 (≥ 2) doses of standard preparation denoted by $s_1, s_2, ..., s_{m_1}$ and m_2 (≥ 2) doses of test preparation denoted by $t_1, t_2, ..., t_{m_2}$. A parallel line assay is called symmetric if $m_1 = m_2$ and asymmetric otherwise. These doses are equi-spaced on logarithmic scale, the common ratio being the same for both the preparations. We thus have $s_i = c_1 h^{i-1}\ (1 \leq i \leq m_1),\ t_i = c_2 h^{i-1}\ (1 \leq i \leq m_2)$, where c_1, c_2 and h (>1) are positive constants. Let τ_i denote the effect of i^{th} dose of standard preparation, $1 \leq i \leq m_1$ and for $1 \leq i \leq m_2$, τ_{m_1+i} be that of i^{th} dose of test preparation. Then, the three contrasts, namely *preparation, combined regression* and *parallelism* respectively given by

$$\psi_p = \left(m_1^{-1}\mathbf{1}'_{m_1}, -m_2^{-1}\mathbf{1}'_{m_2}\right)\boldsymbol{\tau}, \quad \Psi_1 = \delta_1(\mathbf{w}'_1, \mathbf{w}'_2)\boldsymbol{\tau}, \quad \psi'_1 = \delta_2(\theta_2\mathbf{w}'_1, -\theta_1\mathbf{w}'_2)\boldsymbol{\tau}$$

where $\boldsymbol{\tau} = (\tau_1, \tau_2, ..., \tau_v)'$ is the vector of dose effects. $v = m_1 + m_2$, $\mathbf{1}_s$ is a $s \times 1$ vector of all unities, $\delta_1 = 12/\{(\theta_1 + \theta_2)\log h\}, \delta_2 = 12/\{\vartheta_1\theta_2 \log h\}$ and for $i = 1, 2$, $\mathbf{w}'_i = (1, 2, ..., m_i)' - \frac{1}{2}(m_i + 1)\mathbf{1}'_{m_i}$. and for $i = 1, 2$, $\theta_i = m_i(m_i^2 - 1)$.

2.1.2Analysis of parallel line assays

The purpose of analysis of variance of parallel line assays is two fold viz., (i) to test through ANOVA whether the dose metameter and response relationship is linear and (ii) the two lines for the two preparations are parallel. If the tests reveal that the relationship is linear and the lines are parallel, then the relative potency of the test preparation can be estimated. The analysis of the data obtained from a parallel line assay can be carried out as per procedure of the design adopted. LetThe preparation contrast, combined regression contrast and parallelism contrast then can be estimated and significance can be tested as per procedure of contrast analysis. Let there be m_1 ($\geq$2) doses of standard preparation denoted by $s_1, s_2, ..., s_{m_1}$ and m_2 ($\geq$2) doses of test preparation denoted by $t_1, t_2, ..., t_{m_2}$. The total number of treatments (doses) are $m_1 + m_2 = v$ with r_i as the replication number for treatment *i,* $1 \leq i \leq v$. If the design adopted is a completely randomized design, then an outline of analysis of variance is

Table 2.1.1: Analysis of Variance in $m_1 + m_2$ point assays for validity tests:CRD

Source of Variation	d.f.	SS	MS	F
Doses	v-1 = $(m_1 + m_2 - 1)$	SST		
Preparation(Ψ_p)	1	SSLP		
Combined regression (Ψ_1)	1	SSCR		
Parallelism (Ψ_1')	1	SSP	$s_b^2 = SSP/1$	s_b^2 / s^2
Deviation from regression	v - 4	SSDR: by subtraction	$s_d^2 = SSDR/(v-4)$	s_d^2 / s^2
Within doses(error)	$\sum_{i=1}^{v} r_i - v$	SSE	$s^2 = SSE/(\sum_{i=1}^{v} r_i - v)$	
Total	$\sum_{i=1}^{v} r_i - 1$	TSS		

If the design is symmetric parallel line assays with 2*m* doses each replicated *r* times, then the outline of Analysis of variance simplifies to

Source of Variation	**d.f.**	**SS**	**MS**	**F**
Doses	$2m$-1	SST		
Preparation(Ψ_p)	1	SSLP		
Combined regression (Ψ_1)	1	SSCR		
Parallelism (Ψ_1')	1	SSP	$s_b^2 = SSP/1$	s_b^2 / s^2
Deviation from regression	$2m$-4	SSDR: by subtraction	$s_d^2 = SSDR/(2m-4)$	s_d^2 / s^2
Within doses(error)	$2m(r$-1)	SSE	$s^2 = SSE/(2m(r-1))$	
Total	$2rm$-1	TSS		

If the experiment is conducted using a block design with $n = \sum_{i=1}^{v} r_i$ experimental units arranged in b blocks with same or different sizes, then an outline of analysis of variance table for a block design is

Table 2.1.2: Analysis of Variance in $m_1 + m_2$ point assays for validity tests:Block Design

Source of Variation	**d.f.**	**SS**	**MS**	**F**
Between Blocks	b-1	SSB		
Doses (adjusted)	v-1= ($m_1 + m_2 - 1$)	SST		
Preparation(Ψ_p)	1	SSLP		
Combined regression (Ψ_1)	1	SSCR		
Parallelism (Ψ_1')	1	SSP	$s_b^2 = SSP/1$	s_b^2 / s^2
Deviation from regression	$v-4$	SSDR: by subtraction	$s_d^2 = SSDR/(v-4)$	s_d^2 / s^2
Within doses(error)	$\sum_{i=1}^{v} r_i - v$	SSE	$s^2 = SSE/(\sum_{i=1}^{v} r_i - v)$	
Total	$\sum_{i=1}^{v} r_i - 1$	TSS		

For testing the linearity of regression, the mean squares for the deviations from regression is tested by the *F*-test using the within squares as error. For testing parallelism, the "parallelism" component is tested. If both these are not significant, then the relative potency can be estimated.

Example 2.1.1: (Finney, 1978, pp139): This is an example related to the assay of a test preparation of the testosterone propionate against a standard, using three doses of each. Each of the six doses was injected into five capons, and the birds responded by showing a growth of comb. The experiment was conducted using a completely randomized design. The response used for bioassay is the increase in the sum of the length and height of the comb. The data obtained is given below:

c1=c2=20,h=2

Doses→ Responses↓	**Standard Preparation**			**Test Preparation**		
	20µg	**40µg**	**80µg**	**20µg**	**40µg**	**80µg**
	20.2^0	**20.2^1**	**20.2^2**	**20.2^0**	**20.2^1**	**20.2^2**
1	6	12	19	6	12	16
2	6	11	14	6	11	18
3	5	12	14	6	12	19
4	6	10	15	7	12	16
5	7	7	14	4	10	15

For the 6-point assay, the contrasts are

Contrast	s_1	s_2	s_3	t_1	t_2	t_3	**Coefficient Divisor**
	20.2^0	**20.2^1**	**20.2^2**	**20.2^0**	**20.2^1**	**20.2^2**	
Preparation (Ψ_p)	1	1	1	-1	-1	-1	3
Combined Regression (Ψ_1)	-1	0	1	-1	0	1	$6/(\theta \log h)$ =6/(3*(3^2-1)*$\log_{10}2$) =1.204
Parallelism contrast (Ψ_1')	-1	0	1	1	0	-1	$12/(\theta \log h)$ =12/(3*(3^2-1)*$\log_{10}2$) =0.602

The ANOVA table for the above is

Source of variation	**d.f.**	**SS**	**MS**	**F**	**Prob>F**
Doses	5	519.067	103.813	43.86	0.0001
Preparation (ψ_p)	1	4.800	4.800	2.03	0.1673
Combined regression (ψ_1)	1	510.050	510.050	215.51	0.0001
Parallelism (ψ_1')	1	4.050	4.050	1.71	0.2032
Deviation from regression	2	0.167	0.835	<1	NS
Within doses (error)	24	56.800	2.367		
Total	29	575.867			

We can see that both assumptions of linearity of regression and parallelism hold. Therefore, one has to obtain the estimates of the preparation and combined regression contrasts. The estimated values of these contrasts are

Contrast	Estimate	SE of estimate	T for H_0: Contrast =0	Prob>T
Preparation	-0.800	0.562	-1.42	0.1673
Combined Regression	16.777	1.143	14.68	0.0001
Parallelism	-2.990	2.286	-1.31	0.2032

Therefore, the estimate of relative potency is

$$\hat{\rho} = (c_1/c_2)\exp(-\Psi_p/\Psi_1) = \frac{20}{20}\exp(0.800/16.777) = 1.048\,.$$

Example 2.1.2 (Das and Giri (1986), pp294): Here, the data obtained from a 6 point symmetrical parallel line assay collected on a vitamin D assay by Coward and Kassner (1941) has been used with some modifications. The design used is a randomized complete block designs with liters as blocks (12 litters). To ensure comparability of the estimate of relative potency all observation were used, but were fitted into an incomplete block design of the present series by omitting two observations from each of the original block (litters), as shown by blanks in Table 4, and forming 6 additional blocks (13-18) from the 24 observations omitted, ignoring litter differences, but retaining the dose-observation relations. The design assumed is that for 6 treatments in blocks of size 4. The data and the assumed design are shown in the following table

Doses	Standard Preparation			Test Preparation			Block Totals
	s_1 2.5	s_2 5	s_3 10	t_1 2.5	t_2 5	t_3 10	
Blocks							
1	2	8	-	-	9	7	26
2	6	-	9	3	-	8	26
3	-	6	12	4	6	-	28
4	9	11	-	-	14	13	47
5	10	-	17	8	-	10	45
6	-	7	5	-	6	9	27
7	4	10	-	11	13	-	38
8	11	-	9	3	-	15	38
9	-	9	14	5	8	-	36
10	4	7	-	10	10	-	31
11	12	-	9	15	-	15	51
12	-	8	11	-	7	8	34
13	4	4	-	-	5	9	22
14	7	-	8	3	-	9	27
15	-	15	10	6	8	-	39
16	2	4	-	-	6	6	18
17	4	-	13	5	-	12	34
18	-	10	13	4	18	-	45
Dose Total	**75**	**99**	**130**	**69**	**112**	**127**	**612**

The ANOVA table for this example is given by

Nature of Variation	**d.f.**	**s.s**	**m.s.**	**F**	**Prob.>F**
Between Blocks (unadjusted)	17	358.00	21.06	-	-
Doses (adjusted)	5	302.333	60.47	8.83	<0.0001
Between Blocks (adjusted)	17	382.333	22.49	3.28	0.0006
Preparation	1	0.222	0.22	0.03	0.8578
Regression	1	266.02	266.02	38.83	<0.0001
Parallelism	1	7.563	7.563	1.10	0.2986
Deviation from regression	2	28.428	14.214	2.075	0.1370
Error (by subtraction)	49	335.667	6.850		
Total	**71**	**996.00**			

We can see that both assumptions of linearity of regression and parallelism hold.

One has to obtain the estimates of the preparation and combined regression contrasts. The estimated values of these contrasts is

Contrast	Estimate	SE of Estimate	T for H_0: Contrast=0	Prob.>T
Preparation	-0.111	0.617	-0.18	0.8578
Combined regression	7.821	1.255	6.23	<0.0001
Parallelism	-3.045	2.898	-1.05	0.2986

Therefore, the estimate of relative potency is

$$\hat{\rho} = (c_1/c_2)\exp(-\Psi_p/\Psi_1) = \frac{2.5}{2.5}\exp(0.111/7.821).$$

2.2 Slope Ratio Assays

In slope ratio assays, the equations of regression lines for the two preparations in a slope ratio are given by

$$E(y_s) = \alpha + \beta_s x_s \text{ and } E(y_t) = \alpha + \beta_t x_t$$

where $y_s\,(y_t)$ is response due to standard (test) preparation ; $\beta_t = \beta_s \rho^{\lambda}$ and λ is a known constant determined by the transformation. That is, it is assumed that the two regression lines intersect at the same point on the response axis, *i.e.,* the lines have the same intercept. Since the dose takes value zero at response axis, it is logical to include a *blank dose* in the assay for validity test. If there are $m_1(m_2)$ doses in standard (test) preparation then a slope ratio assay contains $m_1 + m_2 + 1$ doses.

The regression coefficients β_s and β_t of the two lines can be estimated subject to the fact that the two lines intersect on the response axis by fitting the relation:

$E(y) = \alpha + \beta_s x_s + \beta_t x_t$ (Unblocked set up)

and $E(y) = \alpha + \beta_s x_s + \beta_t x_t +$ block effect (block design).

Inclusion of blank dose in the assay raises a question "does the linearity of the relation holds up to zero dose?" It is therefore, necessary to test if the relation $E(y) = \alpha + \beta_s x_s + \beta_t x_t$ holds up to zero dose. This provides one validity test. For this, the difference between $\bar{\alpha}$ and $\bar{\alpha}'$ is tested, where, $\bar{\alpha}$ is the estimate of α when the relation is fitted by including the blank dose and $\bar{\alpha}'$ is the estimate of α when the relation is fitted by omitting the observations from the blank dose. This contrast is known as *blank contrast.*

The next question is "Whether the two lines intersect on the response axis?". For this the two lines are fitted individually ignoring the blank dose and then their intercepts on the response axis are compared. This provides another validity test. The corresponding contrast is called *intersection contrast*. There are therefore, two major contrasts of interest in slope ratio assays viz., *blank contrast* and *intersection contrast.*

For slope ratio assay with $(m_1 + m_2 + 1)$ doses, the *blank* and *intersection* contrasts are given by

Blank contrast: $\mathrm{L_B} = (\mathrm{g}, \boldsymbol{\alpha}_1', \boldsymbol{\alpha}_2')\boldsymbol{\tau}$,

Intersection Contrast: $\mathrm{L_I} = \left(0, -\boldsymbol{\alpha}_1^{*'}, \boldsymbol{\alpha}_2^{*'}\right)\boldsymbol{\tau}$,

Where $\boldsymbol{\alpha}_\mathrm{i} = (2 - 2\mathrm{m_i}, 5 - 2\mathrm{m_i}, ..., \mathrm{m_i} - 1)', \mathrm{i} = 1,2;$

$\boldsymbol{\alpha}_\mathrm{i}^* = \dfrac{2}{\mathrm{h_i}}\boldsymbol{\alpha}_\mathrm{i}, \mathrm{i} = 1,2;\ \ \boldsymbol{\tau} = \left(\tau_0, \tau_1, ..., \tau_{\mathrm{m_1}}, \tau_{\mathrm{m_1}+1}, ..., \tau_{\mathrm{m_1}+\mathrm{m_2}}\right)'$ and

$\mathrm{g} = \dfrac{1}{2}\sum_{\mathrm{i}=1}^{2} \mathrm{h_i}$, $\mathrm{h_i} = \mathrm{m_i}(\mathrm{m_i} - 1)$; i=1, 2.

For symmetric slope ratio assay i.e., for $m_1 = m_2 = m$ these contrasts are given by

$\mathrm{L_B^{sy}} = (\mathrm{g_1}, \boldsymbol{\alpha}', \boldsymbol{\alpha}')\boldsymbol{\tau}; \qquad \mathrm{L_I^{sy}} = (0, -\boldsymbol{\alpha}', \boldsymbol{\alpha}')\boldsymbol{\tau}$

where

$\mathrm{g_1} = \mathrm{m}(\mathrm{m} - 1);\ \boldsymbol{\alpha} = (2 - 2\mathrm{m}, 5 - 2\mathrm{m}, ..., \mathrm{m} - 1)';\ \boldsymbol{\tau} = (\tau_0, \tau_1, \tau_2, .., \tau_{2\mathrm{m}})$

Here, the superscript *sy* stands for symmetric.

Now using the above contrasts, one can test the validity and then by fitting $\mathrm{E(y)} = \alpha + \beta_\mathrm{s}\mathrm{x_s} + \beta_\mathrm{t}\mathrm{x_t}$ can obtain the estimate of relative potency.

NESTED BLOCK DESIGNS

Rajender Parsad
I.A.S.R.I., Library Avenue, New Delhi – 110 012
rajender@iasri.res.in

1. Introduction

Heterogeneity in the experimental material is the most important problem to be reckoned within the statistical design of scientific experiments. As if it is not suitably taken care of, in designing of an experiment, is likely to over shadow the real treatment differences making them remain undetected, unless they are large enough. Occasionally, one can find a certain factor (called nuisance factor) which, though not of interest to the experimenter, does contribute significantly to the variability in the experimental material. Various levels of this factor are used for blocking. Blocking is the technique used to bring about homogeneity of experimental units within a block, so that the treatment contrasts are estimated, making use of the intra-block information, with higher efficiency. For the experimental situations where there is only one nuisance factor, the block designs are used. When two such cross-classified factors are present row-column designs such as latin square, lattice square, Youden square, generalized Youden, pseudo Youden designs etc. are being used. In many a field and laboratory experiments the experimental units or conditions differ due to several factors which influence the response under study. It might not always be possible to remove such heterogeneity in response due to the factors other than treatments by blocking alone. There are the experimental situations in which there are one or more factors nested within the blocking factor. When there is one such factor, then nested block designs have been developed. To make the idea more clear, let us have a look at the following experimental situations.

Experimental Situation 1.1: This example relates to a virological experiment, quoted by Preece (1967). Suppose the half-leaves of a plant form the experimental units, on which a number of treatments, say, inoculations with sap from tobacco plants infected with Tobacco Necrosis virus, are to be applied. Suppose the number of treatments is more than the number of suitable half-leaves per plant. Now, there is one source of variation present due to the variability among plants. Further, leaves within a plant may exhibit variation between themselves due to their being located on the upper branch, middle branch or on the lower branch of the same plant. Thus, leaves within plants form a nested 'nuisance' factor, the nested being within the plants. The half-leaves being experimental units, we then have two systems of 'blocks', leaves (which may be called sub-blocks), being nested within plants (which may be called blocks).

Experimental Situation 1.2: In experiments with animals, generally littermates (animals born in the same litter) are experimental units within a block, *i.e.* litters are blocks. However, animals within the same litter may be varying in their initial body weight. If body weight is taken as another blocking factor, we have a system of nested blocks within a block.

Experimental Situation 1.3: Consider a field experiment conducted using a block design and harvesting is done blockwise. To meet the objectives of the experiment, the harvested

samples are to be analysed for their contents in the laboratory by different technicians at same time or by a technician over different periods of time. Therefore, to control the variation due to technicians or time periods, this is taken as another blocking factor, we have a system of nested (sub) blocks *i.e* technicians or time periods within a block.

Hence, for the experimental situations described above, we have one universe for which the results of the experiment will be valid. Out of this universe b_1 blocks of size k_1 have been selected and within each block, there are m sub-blocks such that sub-block size $k_2 = k_1/m$ and total number of the experimental units required is $b_1k_1 = b_1mk_2$.

2. Nested Balanced Incomplete Block Designs

Kleczkouski (1960) devised a form of nested incomplete block design with $v = 8$ treatments for a series of experiments in which bean plants, in the two primary leaves stage, were inoculated with sap from tobacco plants infected with tobacco necrosis virus. The treatments were eight different virus concentrations. Each leaf had two inoculations, one for each half-leaf. Ignoring the leaf positions, plants and leaves were, respectively, the blocks (of size 4) and sub-blocks (of size 2) of a nested balanced incomplete block designs. Preece (1967) has for the case of two-way elimination of heterogeneity, one nested within the other, introduced a Nested Balanced Incomplete Block (NBIB) design.

Definition 2.1: An arrangement of v treatments each replicated r times in two system of blocks is said to be a NBIB design with parameters ($r, v, b_1, k_1, \lambda_1, b_2, k_2, \lambda_2, m$) if

(a) the second system is nested within the first, with each block from the first system (block) containing exactly m blocks from the second system (sub-blocks);
(b) ignoring the second system leaves a balanced incomplete block (BIB) design with b_1 blocks each of k_1 units, with λ_1 concurrences;
(c) ignoring the first system leaves a BIB design with $b_2 = b_1m$ blocks each of $k_2 = k_1/m$ units with , λ_2 concurrences.

Parametric Relationship

1. $vr = b_1k_1 = b_1k_2m = b_2k_2$;
2. $\lambda_1(v-1) = r(k_1-1)$; $(v-1)\lambda_2 = r(k_2-1)$ and
3. $(\lambda_1-m,\lambda_2)(v-1) = r(m-1)$.

Preece (1967) presented the analysis of these designs and provided a list of designs with $r \leq 15$. Jimbo and Kuriki (1983), Dey, Das and Banerjee (1986), Parsad, Gupta and Srivastava (1999) and Morgan, Preece and Rees (2001) gave methods of construction of NBIB designs. Morgan, Preece and Rees (2001) prepared an exhaustive catalogue of NBIB designs with $v \leq 16$ and $r \leq 30$. Some of the methods of constructions of NBIB designs are presented in the sequel.

3. Construction of NBIB Design

Method 2.1: Let their exists a BIB design D_1 with parameters $v', b', r', k' = s^2, \lambda'$, where s is a prime or prime power. Consider another resolvable BIB design D_2 with parameters $v'' = s^2, b'' = s(s+1), k'' = s, r'' = s+1, \lambda'' = 1$. Now, take j^{th} block contents of D_1 as treatments and write a BIB design D_2 in these treatments and arrange the blocks of D_2, replication wise.

Repeat this process for all blocks of D_1. This process results into a NBIB design with parameters with group of blocks of D_2 forming a complete replicate is as block and blocks of D_2 within replications as sub-blocks.

$v = v', r = r'(s+1), b_1 = (s+1)b', k_1 = s^2, b_2 = s(s+1)b', k_2 = s,$
$\lambda_1 = (s+1)\lambda', \lambda_2 = \lambda'.$ (3.1)

Example 3.1: Let D_1 be a BIB design ($v' = 7 = b', r' = 4 = k' = 2^2, \lambda' = 2$), a solution of which can be obtained by developing the initial block (3,5,6,7) mod 13. Let D_2 be the resolvable BIB design ($v'' = 4 = 2^2, b'' = 6 = 2(2+1), r'' = 3 = 2+1, k'' = 2 = s, \lambda'' = 1$). The contents of D_2 are

Replication – I		**Replication - II**		**Replication - III**	
Block - 1	**Block-2**	**Block-3**	**Block-4**	**Block-5**	**Block-6**
A	C	B	A	C	B
B	D	C	D	A	D

Then following the procedure of Method 2.1 *i.e.* by writing D_2, replication wise taking block contents of D_1 as treatments we get the design.

[(3,5),(6,7)]	[(5,6),(3,7)]	[(6,3),(5,7)]
[(4,6),(7,1)]	[(6,7),(4,1)]	[(7,4),(6,1)]
[(5,7),(1,2)]	[(7,1),(5,2)]	[(1,5),(7,2)]
[(6,1),(2,3)]	[(1,2),(6,3)]	[(2,6),(1,3)]
[(7,2),(3,4)]	[(2,3),(7,4)]	[(3,7),(2,4)]
[(1,3),(4,5)]	[(3,4),(1,5)]	[(4,1),(3,5)]
[(2,4),(5,6)]	[(4,5),(2,6)]	[(5,2),(4,6)]

The parameters of the above design are:
$v = 7, r = 12, b_1 = 21, k_1 = 4, \lambda_1 = 6, b_2 = 42, k_2 = 2, \lambda_2 = 2.$

Method 3.2: Suppose there exists a BIB design $(v', b', r', k', \lambda')$ for which an initial block solution based on t-initial blocks is available. Suppose it is possible to divide each initial block into m sub-blocks, each of size k_2, such that the mt sub-blocks form the initial block for generating a BIB design with v' treatments and block size k_2. Then, clearly by developing these initial blocks, we get a NBIB design with parameters

$v = v', r = r', b_1 = b', k_1 = k', \lambda_1 = \lambda', b_2 = mt\ v', k_2, \lambda_2 = r'(k'-m)/m(v-1).$ (3.2)

Using this method, several series of NBIB designs have been obtained. For details see Dey, Das and Banerjee (1986).

Example 2.2: We can construct an NBIB designs with parameters $v = 8, r = 14, b_1 = 28, k_1 = 4, \lambda_1 = 6, b_2 = 56, k_2 = 2, \lambda_2 = 2$. The initial blocks for the design are

[(1,5), (2,3)], [1,6), (4,7)], [(3,5), (8,6)], [(2,1), (8,4)].

We get the design by developing these blocks mod 7 and keeping fix the treatment 8.

Corollary 3.2.1: Following the procedure of Method 2.2, it can easily be seen that following series of NBIB designs can always be constructed.

Series I: $v = 2t+1 = b_1 = 2t+1,\ b_2 = t(2t+1),\ k_1 = 2t,\ k_2 = 2,\ r = 2t,\ \lambda_1 = 2t-1,\ \lambda_2 = 1.$

Series II: $v = 2t,\ b_1 = 2t-1,\ b_2 = t(2t-1),\ k_1 = 2t,\ k_2 = 2,\ r = (2t-1),\ \lambda_1 = 2t-1\ \lambda_2 = 1$

Series I: It is obtained by developing the initial blocks
[*(1,v-1) (2,v-2), ..., (t,v-t)*] mod *2t+1*

For *v = 7*
[(1,6), (2,.5),(3,4)]
[(2,7), (3,6), (4,5)]
[(3,1), (4,7), (5,6)]
[(4,2), (5,1), (6,7)]
[(5,3), (6,2), (7,1)]
[(6,4), (7,3), (1,2)]
[(7,5), (1,4), (2,3)].

Series II: It is obtained by developing the initial block

[*(1,v) (2,v-1), ..., (t,v-t+1)*] mod (*2t-1*) by taking v-th treatment as invariant

For *v = 6*
[(1,6), (2,.5),(3,4)]
[(2,6), (3,1), (4,5)]
[(3,6), (4,2), (5,1)]
[(4,6), (5,3), (1,2)]
[(5,6), (1,4), (2,3)].

In series II, it can easily be seen that the blocks form an RCB design and sub-blocks for BIB designs. Such designs can be treated as complete blocks and incomplete sub-block designs.

Method 3.3: All 1-resolvable BIB designs can be treated as NBIB designs with complete blocks and incomplete sub-blocks, by considering 1-complete replication as blocks and blocks within replication as sub-blocks.

Example 3.3: For an resolvable BIB design v = 9, b = 12, r = 4, k = 3, λ = 1, we get NBIB design with parameters $v = 9,\ b_1 = 4,\ b_2 = 12,\ r = 4,\ k_1 = 9,\ k_2 = 3,\ \lambda_1 = 4,\ \lambda_2 = 1, m = 3.$

[(1,2,3), (4,5,6), (7,8,9)]
[(1,4,7), (2,5,8), (3,6,9)]
[(1,6,8), (2,4,9), (3,5,7)]
[(1,5,9), (2,6,7), (3,4,8)]

The above method can also easily be applied to α-resolvable BIB design.

4. Analysis of Nested Block Designs

Suppose that v treatments are arranged in a equireplicated proper nested block design involving b_1 blocks each of size k_1 and there being m mutually exclusive and exhaustive sub-blocks of size k_2 within the j^{th} *block, j =1(1)* b_1, so that $b_2 = mb_1$ is the total number of sub-blocks. Let $\mathbf{N} = ((n_{ij}))$ be the $v \times b_1$ treatments-blocks incidence matrix, where n_{ij} denotes the number of replications of the i^{th} treatment in the j^{th} block, *i =1(1)v*. The row sums of $\mathbf{N}$ are $r\mathbf{1}_v$ and column sums are $k_1\mathbf{1}_{b_1}$ where r and k_1 denote respectively the replication number of

treatments and block sizes. Let $\mathbf{M}=((m_{ij'(j)}))$ denote the $v \times b_2$ treatments-sub-blocks matrix, where $m_{ij'(j)}$ denotes the replication number of i^{th} treatment in the j'^{th} sub-block nested with in the j^{th} block. The row sums of $\mathbf{M}$ are $r\mathbf{1}_v$ while its column sums are $k_2\mathbf{1}'_{b_2}$. Here k_2 denotes the size of j'^{th} sub-block nested within the j^{th} block. Let $\mathbf{R} = r\,\mathbf{I}$, $\mathbf{K}_1 = k_1\mathbf{I}$ and $\mathbf{K}_2 = k_2\,\mathbf{I}$.

The model considered for the analysis of these designs is

$$y_{ij'(j)u} = \mu + \tau_i + \beta_j + \eta_{j'(j)} + e_{ij'(j)u} \qquad (2.3)$$

where $y_{ij'(j)u}$ is the u^{th} observation obtained from the i^{th} treatment in the j'^{th} sub-block, $u = 1,...,m_{ij'(j)}$, *i.e.* the number of replications of i^{th} treatment in j'^{th} sub-block, $j' = 1, ..., m$ nested within the j^{th} block, $j = 1(1)b_1$. μ is the general mean, τ_i is the i^{th} treatment effect, β_j is the j^{th} block effect, $n_{ij'(j)}$ is the effect of the j'^{th} sub-block nested within the j^{th} block, and the quantities $e_{ij'(j)u}$ are uncorrelated errors with mean zero and common variance σ^2.

Let $\mathbf{T} = (\mathbf{T}_1,...,\mathbf{T}_v)'$ be the $v \times 1$ vector of treatment totals and $\mathbf{V} = (\mathbf{V}_1,...,\mathbf{V}_{b_2})'$ be the $b_2 \times 1$ vector of sub-block totals, $\mathbf{B} = (\mathbf{B}_1,...,\mathbf{B}_{b_1})'$ be the $b_1 \times 1$ vector of block totals G be the grand total. Then under fixed effects model (1.1), using the principle of least squares, the reduced normal equations for obtaining best linear unbiased estimates of treatment contrasts is given by

$$\mathbf{C}\tau = \mathbf{Q} \qquad (2.4)$$

where $\mathbf{C} = \mathbf{R} - \mathbf{M}\,\mathbf{K}_2^{-1}\,\mathbf{M}'$ and $\mathbf{Q} = \mathbf{T} - \mathbf{M}\mathbf{K}_2^{-1}\mathbf{V}$

such that $Q_i = T_i - \sum_{j'(i)=1} m_{ij'(j)} V_{j'} / k_{2j'}$.

The above information matrix is same as the usual **C -** matrix that is obtained if blocks are ignored and the design is analysed treating sub-blocks as blocks. Therefore, in so far as the estimation of treatment effects is considered, it is only the sub-block structure that matters. The vxv matrix **C** is symmetric, non-negative definite with zero row-sums. For a connected nested block design, the Rank (**C**) = *v-1*. A solution of (2.4) is $\hat{\boldsymbol{\tau}} = \mathbf{C}^{-}\mathbf{Q}$ and treatment sum of squares (eliminating blocks is given by $\mathbf{Q}\mathbf{C}^{-}\mathbf{Q}$. For a NBIB design $\mathbf{C} = \frac{\lambda_2 v}{k_2}(\mathbf{I} - \frac{1}{v}\mathbf{1}\mathbf{1}')$ and $\mathbf{C}^{-} = \frac{k_2}{\lambda_2 v}\mathbf{I}$. Therefore, treatment sum of squares for a NBIB design (SST) is $= \frac{k_2}{\lambda_2 v}\sum_{i=1}^{k_2} Q_i^2$. Under a fixed effects model (1.1) the outline of ANOVA is given as below:

ANOVA

Source of Variation	**D.F.**	**S.S.**

Between Blocks	b_1-1	$\sum_{j=1}^{b_1}\frac{B_j^2}{k_1}-\frac{G^2}{n}$
Between Sub-blocks within blocks	$b_1(m-1)=b_2-b_1$	$\sum_{j'=1}^{b_2}\frac{V_{j'}^2}{k_2}-\sum_{j=1}^{b_1}\frac{B_j^2}{k_1}$
Treatments	v-1	$\frac{k_2}{\lambda_2 v}\sum_{i=1}^{k_2}Q_i^2$
Error	$n- b_1m-v+1$	By subtraction
Total	n-1	$\sum_i\sum_j\sum_{j'}y_{ij'(j)}^2-\frac{G^2}{n}$

As in the analysis of a balanced incomplete block design, the sum of squares for blocks (eliminating treatments) splits into parts, one affects by treatment differences, the other pure intra-block error. Similarly, in the analysis of a NBIB design, the sums of squares for blocks (eliminating treatments) and sub-blocks within blocks (eliminating treatments) each split into a treatment component and a pure error component. The bifurcation of various d.f is given in Table 1.

Table - 1

Source of Variation	**D.F.**
Blocks (ignoring treatments)	
Treatment Component	v-1
Remainder	b_1-v
Total	b_1-1
Sub-blocks within blocks(ignoring treatments)	
Treatment Component	v-1
Remainder	b_2-b_1- v+1
Total	b_2-b_1
Treatments(eliminating sub-blocks)	v-1
Intra-sub-block error	rv-v-b_2+1
Total	rv-1

Using a mixed effects model with block and sub-block effects as random, three independent estimates of treatment contrast *viz.*(i) within sub-block estimate, (ii) between sub-blocks within block estimate and (iii) between block estimate are obtained. These are then combined linearly, to obtain most efficient estimates of treatment contrasts or differences for further analysis and inferences.

Gupta (1993) showed that the information matrix for estimating the treatments effects of nested block design is same as that of information matrix of a block design considering sub-blocks as blocks and this result hold both for proper and non-proper settings.

A connected nested block design is variance balanced iff,

$$\mathbf{C} = \theta(\mathbf{I}_v - \mathbf{I}_v \mathbf{1}\mathbf{1}') \qquad (2.5)$$

where $\theta = \frac{n - b_2}{v - 1}$, a scalar constant, is the unique positive eigen value of **C** with multiplicity *v-1* and $\mathbf{I}_v$ is an identity matrix of order *v*, $\mathbf{I}_v$ is a $v \times 1$ vector of unities.

Gupta (1993) also showed that optimality results of nested block designs are dependent on the optimality results of sub-blocks designs. He also obtained several series of optimal nested block designs with unequal block sizes. As a particular case of these results, it can easily be verified that a NBIB design whenever existent is universally optimal over the competing class of designs.

For an excellent review of nested block designs and their generalization one may refer to Calinski and Kageyama (1996).

5. Nested Partially Balanced Incomplete Block designs

A nested variance balanced block design may not exist always or even if it exists may require a large number of replications which the experimenter may not be able to afford. To deal with such situations, Homel and Robinson (1975) introduced nested partially balanced incomplete block (NPBIB) designs as defined below:

Definition 1.1: An NPBIB design based on *m* (≥ 2)-class association scheme defined in *v* symbols, is an arrangement of *v* symbols into b_2 sub-blocks of size k_2 nested within b_1 (= b_2/t, *t* is an integer) blocks of size k_1 (= $tk_2 < v$) such that

(i) Every symbol occurs at most once in a block.

(ii) Every symbol appears at most *r* times in the design.

(iii) If two symbols, say α and β, are i^{th} associates, then they occur together in λ_{1i} blocks and λ_{2i} sub-blocks, the numbers λ_{1i}, λ_{2i} being independent of the particular pair of i^{th} associates α and β, $i = 1, 2 \ldots m$.

The numbers $v, b_1, b_2, r, k_1, k_2, \lambda_{1i}, \lambda_{2i}$ ($i = 1, 2 \ldots m$) are called parameters of the design. If $\lambda_{1i} = \lambda_1$ and $\lambda_{2i} = \lambda_2$; $\forall\, i = 1, 2 \ldots m$, then an NPBIB design reduces to NBIB design. Since then several methods of construction of NPBIB designs have been obtained and are available in the literature. Satpati and Parsad (2003) have prepared catalogues of two and three associate class NPBIB designs for $v \leq 30$ and $r \leq 15$.

For several factors nested within each other nested multiway designs have been obtained. Optimality aspects of these designs have been studied by Bagchi (1991). NBIB and NPBIB designs with sub-blocks sizes as 2 has an interesting application in obtaining block designs for diallel cross experiments.

References

Bagchi, S.(1991). On optimality of nested multiway designs. *Proceedings of the R.C.Bose Symposium on Probability, Statistics and Design of Experiments:* Delhi 27-30 December, 1988, Wiley Eastern, New Delhi, 23-31.

Calinski, T. and Kageyama, S. (1996). Block Designs : Their combinatorial and statistical properties. *Handbook of Statistics ed. S.Ghosh and C.R.Rao,* **VI 3**, Elsevier Science B.V.

Dey,A., Das, U.S., and Banerjee, A.K. (1986). On Nested balanced incomplete block designs. *Calcutta Statist. Assoc. Bull*.,**35**, 161-167.

Gupta,V.K. (1993). Optimal nested block designs. *J. Indian Soc. Agric***.** *Statist.*, **45(2),** 187-194.

Homel, R.J. and Robinson, J. (1975). Nested partially balanced incomplete block designs. *Sankhy* $\bar{a}$, B37, 201-210.

Jimbo, M. and Kuriki, S. (1983). Constructions of nested designs. *Ars Comb.*, 16, 275-285.

Kleczkouski, A. (1960). Interpreting relationship between the concentrations of plants viruses and number of local lesions. *J.Gen. Microbiol*., **4,** 53-69.

Morgan, J.P., Preece, D.A. and Rees, D.H. (2001). Nested balanced incomplete block designs. *Discrete Mathematics.*, 231, 351-389.

Parsad, R., Gupta, V.K. and Srivastava, R. (1999). Universally optimal block designs for diallel crosses. *Statistics and Applications*, 1, 35-52.

Preece, D.A.(1967).Nested balanced incomplete block designs. *Biometrika*, **54**, 479-486

Satpati, S.K. and Parsad, R. (2005). (2004). Construction and cataloguing of nested partially balanced incomplete block designs. *Ars Combinatoria,* **73**, 299-309.