Experimental designs

Deependra Dhakal

Assistant Professor

Agriculture and Forestry University

https://rookie.rbind.io

- Experiments allow us to prove cause-and-effect relationships
- An experiment involves applying a treatment to a subject, and seeing how it affects the outcome compared to other groups.
- The experimenter must identify at least one explanatory variable (or factor) to manipulate, and at least one response variable to measure.
- The items which receive (or don't receive) the treatments are called experimental units.
  - when people are involved, they are commonly called subjects or participants
- The specific values that the experimenter chooses for a factor are called the levels of the factor.
- A treatment is a combination of specific levels from all the factors that an experimental unit receives.

Example

A farm products manufacturer wants to determine if the yield of a crop is different when the soil is treated with three different types of fertilizer. 15 similar plots of land are planted with the same type of seed but are fertilized differently. At the end of the growing season, the mean yields from the plots will be compared.

- Identify the experimental units.

Example

A farm products manufacturer wants to determine if the yield of a crop is different when the soil is treated with three different types of fertilizer. 15 similar plots of land are planted with the same type of seed but are fertilized differently. At the end of the growing season, the mean yields from the plots will be compared.

- Identify the experimental units.
  - $\longrightarrow$ Plots of land

Example

A farm products manufacturer wants to determine if the yield of a crop is different when the soil is treated with three different types of fertilizer. 15 similar plots of land are planted with the same type of seed but are fertilized differently. At the end of the growing season, the mean yields from the plots will be compared.

- Identify the experimental units.
  - $\longrightarrow$ Plots of land
- Identify the factor(s) in the experiment, and the number of levels for each.

Example

A farm products manufacturer wants to determine if the yield of a crop is different when the soil is treated with three different types of fertilizer. 15 similar plots of land are planted with the same type of seed but are fertilized differently. At the end of the growing season, the mean yields from the plots will be compared.

- Identify the experimental units.
  - $\longrightarrow$ Plots of land
- Identify the factor(s) in the experiment, and the number of levels for each.
  - $\longrightarrow$ Fertilizer (single factor) at 3 levels

Example

A farm products manufacturer wants to determine if the yield of a crop is different when the soil is treated with three different types of fertilizer. 15 similar plots of land are planted with the same type of seed but are fertilized differently. At the end of the growing season, the mean yields from the plots will be compared.

- Identify the experimental units.
  - $\longrightarrow$ Plots of land
- Identify the factor(s) in the experiment, and the number of levels for each.
  - $\longrightarrow$ Fertilizer (single factor) at 3 levels
- Identify the number of treatments.

Example

A farm products manufacturer wants to determine if the yield of a crop is different when the soil is treated with three different types of fertilizer. 15 similar plots of land are planted with the same type of seed but are fertilized differently. At the end of the growing season, the mean yields from the plots will be compared.

- Identify the experimental units.
  - $\longrightarrow$ Plots of land
- Identify the factor(s) in the experiment, and the number of levels for each.
  - $\longrightarrow$ Fertilizer (single factor) at 3 levels
- Identify the number of treatments.
  - $\longrightarrow$ Three

Example

A farm products manufacturer wants to determine if the yield of a crop is different when the soil is treated with three different types of fertilizer. 15 similar plots of land are planted with the same type of seed but are fertilized differently. At the end of the growing season, the mean yields from the plots will be compared.

- Identify the experimental units.
  - $\longrightarrow$ Plots of land
- Identify the factor(s) in the experiment, and the number of levels for each.
  - $\longrightarrow$ Fertilizer (single factor) at 3 levels
- Identify the number of treatments.
  - $\longrightarrow$ Three
- Identify the response variable measured.

Example

A farm products manufacturer wants to determine if the yield of a crop is different when the soil is treated with three different types of fertilizer. 15 similar plots of land are planted with the same type of seed but are fertilized differently. At the end of the growing season, the mean yields from the plots will be compared.

- Identify the experimental units.
  - $\longrightarrow$ Plots of land
- Identify the factor(s) in the experiment, and the number of levels for each.
  - $\longrightarrow$ Fertilizer (single factor) at 3 levels
- Identify the number of treatments.
  - $\longrightarrow$ Three
- Identify the response variable measured.
  - $\longrightarrow$ Mean plot yield

Completely Randomized Design (One way Analysis of Variance)

- Involves selection of random samples from each of k different levels corresponding to the k populations, also the *treatments* for this one-way classification. This design involves only one factor, the population from which the measurement comes – hence the designation as a one-way classification.
- To find out whether the difference exists among the k population means, or not, the analysis of variance procedure provides one overall test to judge the equality of the k population means.

- A researcher is interested in the effects of five types of insecticides for use in controlling the boll weevil in cotton fields. Explain how to implement a completely randomized design to investigate the effects of the five insecticides on crop yield.
  - The only way to generate the equivalent of five random samples from the hypothetical populations corresponding to the five insecticides is to use a method called a randomized assignment.
  - A fixed number of cotton plants are chosen for treatment, and each is assigned a random number. Suppose that each sample is to have an equal number of measurements. Using a randomization device, you can assign the first n plants chosen to receive insecticide 1, the second n plants to receive insecticide 2, and so on, until all five treatments have been assigned.

Suppose there are k population means, $\mu_1, \mu_2, ..., \mu_k$, based on independent random samples of size $n_1, n_2, ..., n_k$ from normal populations with a common variance σ^2. That is, each of the normal populations has the same shape, but their locations might be different.

Let x_{ij} be the jth measurement $(j = 1, 2, ..., n_i)$ in the ith sample. The analysis of variance procedure begins by considering the total variation in the experiment, which is measured by a quantity called the total sum of square (TSS):

$$\text{Total SS} = \sum(x_i j - \bar{x})^2 = \sum x_{ij}^2 - \frac{(\sum x_{ij})^2}{n}$$

The first part of above expression gives the sample variance of the entire set of $n = n_1 + n_2 + ... + n_k$ measurements. The second part of the calculation is called the correction factor (CF). If we let G represent the grand total of all n obervations, then

$$CF = \frac{(\sum x_{ij})^2}{n} = \frac{G^2}{n}$$

The Total SS is partitioned into two components. The first component, called the sum of squares for treatments (SST), measures the variation among the k sample means:

$$SST = \sum n_i(\bar{x}_i - \bar{x})^2 = \sum \frac{T_i^2}{n_i} - CF$$

where T_i is the total of the observations for treatment i. The second component, called the sum of squares for error (SSE), is used to measure the pooled variation within the k samples:

$$SSE = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + ... + (n_k - 1)s_k^2$$

This formula is a direct extension of the numerator in the formula for the pooled estimate of σ^2.

Pooled estimate of variance

The population variance σ^2 describes the shape of the normal distributions from which your samples come, so that either s_1^2 or s_2^2 would give you an estimate of σ^2 . But why use just one when information is provided by both? A better procedure is to combine the information in both sample variances using a weighted average, in which the weights are determined by the relative amount of information (the number of measurements) in each sample. For example, if the first sample contained twice as many measurements as the second, you might consider giving the first sample variance twice as much weight. To achieve this result, use this formula:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

We can show algebraically that, in the analysis of variance,

Total SS = SST + SSE

Therefore, only one of the two sums of squares may be calculated and the third can be found by subtraction.

Each of the sources of variation, when divided by its appropriate degree of freedom, provides an estimate of the variation in the experiment. Since Total SS involves n squared observations, its degree of freedom are $df = (n - 1)$. Similarly, the sum of squares for treatments involves k squared observations, and its degree of freedom are $df = k - 1$. Finally, the sum of squares for error, a direct extension of the pooled estimated, has

$$df = (n_1 - 1) + (n_2 - 1) + ... + (n_k - 1) = n - k$$

Notice that the degrees of freedom for treatments and errors are additive – that is,
$df(\text{total}) = df(\text{treatments}) + df(\text{error})$.

These two sources of variation and their respective degrees of freedom are combined to form the mean square as $MS = SS/df$. The total variation in the experiment is then displayed in an analysis of variance (or ANOVA) table.

Sources of variation	Degree of freedom	Sum of squares (SS)	Mean square (MS)	Expected mean square (MS)
Treatments	$t-1$	$r\sum_i(\bar{y}_i - \bar{y})^2 = SS(T)$	$MS(T) = \frac{SS(T)}{(t-1)}$	$\sigma_e^2 + \frac{r}{t-1}\sum_{i=1}^{t}\tau_i^2$
Error	$t(r-1)$	$\sum_{i,j}(y_{ij} - \bar{y}_i)^2 = SS(E)$	$MS(E) = \frac{SS(E)}{t(r-1)}$	σ_e^2

Randomized Complete Block Design (Two way Analysis of Variance)

Consider an experimental situation in which v treatments are to be compared via $N = vr$ experimental units (plots) arranged in r blocks each of size v such that each treatment occurs exactly once in each block, i.e., the experiment is conducted using a randomized complete block design. Let n plants be selected from each plot and observations are made from n selected plants. The response variable can be represented by a linear, additive, fixed effect model as,

$$Y_{ijt} = \mu + \tau_i + \beta_j + e_{ij} + \eta_{ijt}$$

Where Y_{ijt} is the observation pertaining to the t-th sampling unit for the i-th treatment in the j-th blck ($i = 1, 2, \ldots, v$, $j = 1, 2, \ldots, r$; $t = 1, 2, \ldots, n$), μ is the general mean effect; τ_i is the i-th treatment effect, β_j is the effect of j-th block, e_{ij} is the plot error distributed as $N(0, \sigma_e^2)$, η_{ijt} is the sampling error distributed as $N(0, \sigma_s^2)$.

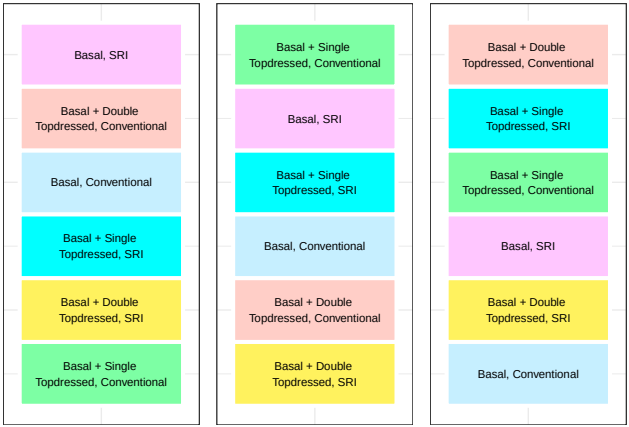


Figure 1: Layout of 2x3 factorial RCBD in randomized experimental layout

The analysis of variance (ANOVA) for such a design is given in 1.

Table 1: ANOVA table for RCB designs

Sources of variation	Degree of freedom	Sum of squares (SS)	Mean square (MS)	Expected mean square (MS)
Blocks	$r-1$	SSB		
Treatments	$v-1$	SST		$\sigma_s^2 + n\sigma_e^2 + \frac{rn}{v-1}\sum_{i=1}^{v}\tau_i^2$
Treatments x Blocks (experimental error)	$(v-1)(r-1)$	SSBT	MSBT	$\sigma_s^2 + n\sigma_e^2$
Sampling error	$rv(n-1)$	SSSE	MSSE	σ_s^2

The sum of squares due to different components of ANOVA can be obtained as follows:

Form a $r \times v$ two-way table between blocks and treatments, each cell figure being the total overall samples from a plot.

Table 2: Two way tabulation of block and treatment observations.

Blocks	1	2	i	v	Block totals
1	T_{11}	T_{21}	T_{i1}	T_{v1}	$B_{.1}$
2	T_{12}	T_{22}	T_{i2}	T_{v2}	$B_{.2}$
.	.	.	.	.	.
j	T_{1j}	T_{2j}	T_{ij}	T_{vj}	$B_{.j}$
.	.	.	.	.	.
r	T_{1r}	T_{2r}	T_{ir}	T_{vr}	$B_{.r}$

The sum of squares (S.S) due to different components of ANOVA can be obtained as follows:

Grand Total (GT) $= \sum_{i=1}^{v} \sum_{j=1}^{r} \sum_{t=1}^{n} y_{ijt}$

Correction factor (CF) $= \frac{GT^2}{rvn}$

Total SS of the table (TSS) $= \frac{\sum_{i=1}^{v} \sum_{j=1}^{r} \sum_{t=1}^{n} y_{ijt}^2}{n} - CF$

T_i = i-th treatment total $= \sum_{j=1}^{r} \sum_{t=1}^{n} y_{ijt}$

B_i = j-th block total $= \sum_{i=1}^{v} \sum_{t=1}^{n} y_{ijt}$

Treatment SS (SST) $= \frac{\sum_{i=1}^{v} T_i^2}{nv} - CF$

Block SS (SSB) $= \frac{\sum_{j=1}^{r} B_j^2}{nv} - CF$

Block x Treatment SS (SSBT) = TSS - SST - SSB

Total SS of the entire data $= \sum_{i=1}^{v} \sum_{j=1}^{r} \sum_{t=1}^{n} y_{ijt}^2 - CF$

Sum of squares due to the sampling error (SSSE) = Total SS of the entire data - SSB - SST - SSBT

Using the expression of expected mean squares in the above ANOVA Table (1), it is clear that the null hypothesis regarding the equality of treatment effects is tested against the experimental error. From the ANOVA, it is also clear that the sampling error is estimated as

$\hat{\sigma_s}^2 = s_2^2$.

The experimental error (variance between plots of the same treatment) is estimated as $\hat{\sigma}_e^2 = \frac{s_1^2 - s_2^2}{n}$. When $\hat{\sigma}_e^2$ is negative, it is taken as zero.

The variance of the i-th treatment mean ($\bar{Y}_i$) based on r-replications and n-samples per plot $= \frac{\sigma_s^2 + n\sigma_e^2}{rn}$

The estimated variance of $\bar{Y}_{i..} = \frac{\hat{\sigma}_s^2 + n\hat{\sigma}_e^2}{rn}$.

Taking the number of sampling units in a plot to be large (infinite), the estimated variance of a treatment mean when there is complete recording (i.e., the entire plot is harvested) $= \frac{\hat{\sigma}_e^2}{r}$

The efficiency of sampling as compared to complete enumeration:

$$\frac{\frac{\hat{\sigma}_e^2}{r}}{\frac{\hat{\sigma}_s^2+n\hat{\sigma}_e^2}{rn}}$$

The standard error of a treatment mean $\bar{Y}_{i...}$ with n samples per plot and r replication is

$$\left[\frac{\hat{\sigma}_s^2}{rn}+\frac{\hat{\sigma}_e^2}{r}\right]^{\frac{1}{2}}$$

The coefficient of variation is

$$p=\frac{\left[\frac{\hat{\sigma}_s^2}{rn}+\frac{\hat{\sigma}_e^2}{r}\right]^{\frac{1}{2}}}{\bar{Y}_{i...}}\times 100$$

Thus, n can be found by re-arranging the above expression.

Generally, the margin of error (d or D) is $Z_{\alpha/2}$ times the value of coefficient of variation of $\bar{Y}_{i\ldots}$ based on the concept of $100(1-\alpha)\%$, confidence intervals. Therefore,

$$n = \frac{\hat{\sigma}_s^2}{r}\left[\frac{Z_{\alpha/2}^2}{D^2(\bar{Y}_i)^2 - Z_{\alpha/2}^2\frac{\hat{\sigma}_e^2}{r}}\right]$$

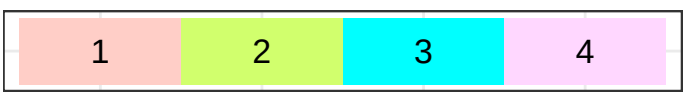
For any given r and $p(D)$, there will be t values for n corresponding to the t treatment means. The maximum n will ensure the estimation of any treatment mean with a standard error not exceeding p percent or margin of error not exceeding D.

For an example to accompany theory of analyzing RCB design, refer to http://apps.iasri.res.in/ebook/EBADAT/2-Basic%20Statistical%20Techniques/22-plotsamp-final.pdf .

Split plot design

Design was developed and first used for agricultural, mainly agronomic experiments.

When we have two treatment factors A and B, with levels $a_1, a_2, ..., a_a$ and $b_1, b_2, ..., b_b$, respectively. Factor A is referred to as the whole-plot-factor and the EUs to which the levels of A are applied are the whole-plots. Factor B is the split-plot factor and the EUs to which the level of B are applied are the split-plots, each whole-plot having b split-plots as illustrated below for $b = 4$



A replicate consists then of one application of each level $a_1, a_2, ..., a_a$ and within each of the a whole-plots of one application of each level $b_1, b_2, ..., b_b$. And the design consists then of r such replications.

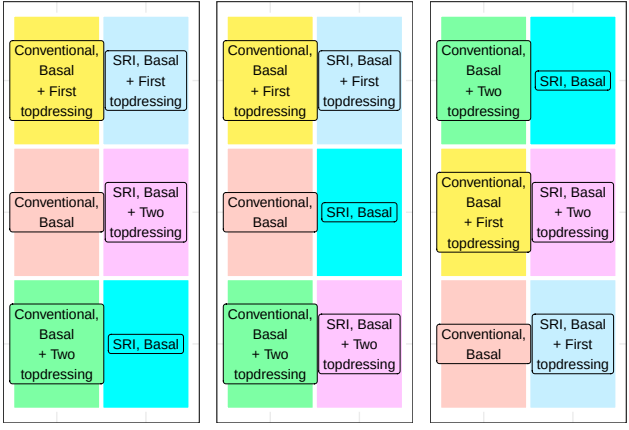


Figure 2: Layout of Split plot design with main (Tillage operation) and sub (Fertilizer application) plot factors in randomized experimental layout

It is useful to think of this arrangement as superimposing one RCBD on top of another RCBD. For the first RCBD, involving the whole-plots and the whole-plot factor, we have

$$\mathrm{RCBD_A} : t = a, \text{number of blocks} = r$$

and for the second RCBD, involving the split-plots and split-plot factor, we have

$$\mathrm{RCBD_B} : t = b, \text{number of blocks} = ra$$

This brings out the fact that two independent randomizations are being used.

Assuming no replicate x B interaction (since we are assuming unit-treatment additivity), we then have the complete partitioning of the d.f. as given in the ANOVA of Table 3.

Table 3: ANOVA table for Split plot designs

Source	df	SS
Replicates	$(r-1)$	$ab\sum_i(\bar{y}_{i..}-\bar{y})^2 = SS(R)$
A-factor	$(a-1)$	$rb\sum_j(\bar{y}_{.j.}-\bar{y})^2 = SS(A)$
Error (A)	$(r-1)(a-1)$	$b\sum_{i,j}(\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y}_{...})^2 = SS(E_A)$
B-factor	$(b-1)$	$ra\sum_k(\bar{y}_{..k}-\bar{y}_{...})^2 = SS(B)$
A x B	$(a-1)(b-1)$	$r\sum_{j,k}(\bar{y}_{.jk}-\bar{y}_{.j.}-\bar{y}_{..k}+\bar{y}_{...})^2 = SS(A\times B)$
Error(B)	$(r-1)a(b-1)$	$\sum_{ijk}(y_{ijk}-\bar{y}_{ij.}-\bar{y}_{.jk}+\bar{y}_{.j.})^2 = SS(E_B)$
Total	$rab-1$	$\sum_{i,j,k}(y_{ijk}-\bar{y}_{...})^2$

Latin Square Design

- Special case of ANOVA model with 3 factors.
  - with the equal number of factor levels in all the independent variables
- A $\nu \times \nu$ latin square is an arrangement of ν Latin letters into a $\nu \times \nu$ array (a table with ν rows and ν columns) in such a way that each letter occurs once in each row and once in each column.

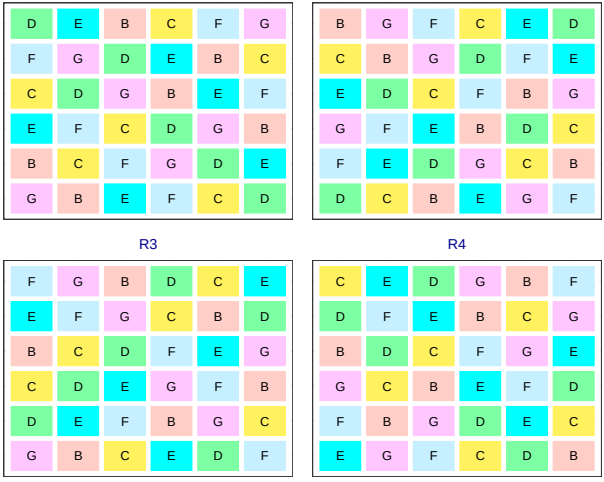


Figure 3: A 6 x 6 Latin square design replicated in 4 blocks.

- A row block or a column block alone, ignoring the other of a Latin square design is analogous to a RCBD.
- Each level of the treatment factor is observed $r = \nu$ times in a single Latin square block, $\therefore$ making replications of the square blocks unnecessary.
  - ♣ less commonly, s $\nu \times \nu$ Latin squares are pieced together (s-replicate Latin square).
- Latin square designs are often used in experiments involving subjects, especially where the subjects are allocated a sequence of treatments over time and where the time effect is thought to have a major effect on the response.

Split-split-plot design

For a 5 x 3 x 3 factorial design, replicated 3 times, ANOVA of split-split-plot is shown in Table 4.

Table 4: ANOVA table for split split plot designs

Source of variation	df	SS	MSS
Mainplot			
Replication (R)	$(r-1) = 3-1 = 2$		
Main-plot factor (A)	$(a-1) = 3-1 = 2$		
Error (a)	$(r-1)(a-1) = 8$		
Subplot			
Subplot factor (B)	$(b-1) = 2$		
A x B	$(a-1)(b-1) = 8$		
Error(b)	$a(r-1)(b-1) = 20$		
Sub-sub plot			
Sub-sub plot factor (C)	$(c-1) = 2$		
A x C	$(a-1)(c-1) = 8$		
B x C	$(b-1)(c-1) = 4$		
A x B x C	$(a-1)(b-1)(c-1) = 16$		
Error (c)	$ab(r-1)(c-1) = 60$		
Total	$rabc - 1 = 134$		

Models	Best used for	Advantages	Disadvantages
Randomized complete block design (RCBD)	Small trials, with seed and space for sufficient replications	Simple to design, simple to analyze, robust to missing plots, broadly accepted	Spatial effects can be problematic in large-sized RCBDs Must have seed for complete replications
Lattice	Trials with many entries, but seed and space for replications	Relatively simple to design Incomplete blocks can better capture field effects, but can resolve to RCBD	Relatively inflexible in numbers of entries, blocks, and reps Has $r(b-1)$ fewer degrees of freedom for error compared to RCBD, where r is the number of reps and b is the number of blocks
α-designs	Trials with many entries, and seed and space for replications	Flexible in terms of number of entries, blocks, and reps Incomplete blocks can better capture field effects, but can resolve to RCBD	More complex to design—generally requires computer assistance to choose best design Has $r(b-1)$ fewer degrees of freedom for error compared to RCBD, where r is the number of reps and b is the number of blocks
Latin square	Trials with few entries that would benefit from two-dimensional blocking	Relatively simple to design Can account for field effects in two dimensions Can resolve to RCBD	Quickly becomes massive, as the number of plots are equal to the square of entries Has $r + c - 2$ fewer degrees of freedom for error compared to RCBD, where r is rows and c is columns
Row–column	Trials with an unknown spatial-effect gradient or with a gradient in two dimensions	Can account for field effects in two dimensions Can be superimposed on other designs such as RCBD and α	Can be complicated to design Creates additional restrictions in degrees of freedom Has $r + c - 2$ fewer degrees of freedom for error compared to RCBD, where r is rows and c is columns
Control plot	Trials with many entries but few replications per entry due to limited seed or space	Relatively simple to design Can account for field variation regardless of direction of spatial gradient	Can be inefficient if a large percentage of the trial field is planted to control variety
Spatial	Trials with expected large and unknown field effects	Can efficiently account for field variation regardless of direction of spatial gradient	Can be complicated to design Spatial term will reduce degrees of freedom for error
Augmented	Trials with many entries but few replications per entry due to limited seed or space	Relatively simple to design Allows more entries to be tested in a smaller space relative to fully replicated designs	Risk of missing entries if plots lost Can have lower power to detect differences between experimental entries

References