RANDOMISED BLOCK DESIGN (R.B.D.)

In field experimentation, if the whole of the experimental area is not homogeneous and the fertility gradient is only in one direction, then a simple method of controlling the variability of the experimental material consists in stratifying or grouping the whole area into relatively homogenous strata or subgroups (or blocks or replicates) perpendicular to the direction of the fertility gradient. Now if the treatments are applied at random into relatively homogeneous units within each strata or block and replicated over all the blocks, the design is a randomized block design (R.B.D.)

In a CRD, we do not resort to grouping of the experimental site and allocate the treatments at random to the experimental units. But in R.B.D. treatments are allocated at random within the units of each stratum or block, i.e., randomization is restricted. Also variation among blocks is removed from variation due to error. Hence, if it is desired to control one source of variation by stratification, the experimenter should select the RBD, rather than CRD.

Applications of R.B.D.

Despite it's agricultural origin, the randomized block design is widely used in many types of studies. For example to determine the differences in productivity of c makes of machines(treatments). We may isolate the possible effects due to differences in efficiency among operators (blocks) by randomly rotating machine assignments in such a way that each operator works on all the machines, the basic idea is to compare treatment levels (the different machines)within block of relatively homogeneous experimental material (the same operator), then repeat the comparison on another block (another operator), and so on for additional repetition of the comparison.

Layout of R.B.D.

In agricultural experiments, the layout of R.B.D, can be given as follows

Let us consider five treatments A, B, C, D and E replicated 4 times. We divide the whole experimental area into four relatively homogenous strata or blocks and each block into five units or plots, treatments are then allocated at random to the plots of a block, fresh randomization done for each block. A particular layout may be as follows:

BLOCK I	A	E	B	D	C
BLOCK II	E	D	C	B	A
BLOCK III	C	B	A	D	E
BLOCK IV	A	D	E	C	B

For randomization we may use Tippet's randomization tables. Let us select one digit numbers in the order of their occurrence leaving 0 and numbers greater than 5. Suppose we get a random numbers as 1,3,5,4 and 2, so in the first block we allocate treatment A to the first plot, C to the second plot, E to the third plot, D to the fourth plot and B to the fifth plot. Similarly, we use fresh random numbers for each of the other three plots and allocate the treatments accordingly.

Advantages of R.B.D.

Chief advantages of R.B.D. are as follows

i) Accuracy: This design has been shown to be more efficient or accurate than C.R.D. for most types of experimental work. The elimination of between sum of squares from residual sum of squares usually results in a decrease of error mean sum of squares.
ii) Flexibility: In R.B.D. no restrictions are placed on the number of treatments or the number of replicates.
iii) Ease of analysis: Statistical analysis is simple, rapid and straight forward.

Disadvantages of R.B.D.

i) R.B.D. may give misleading results if blocks are not homogeneous.
ii) R.B.D. is not suitable for large number of treatments because in that case the block size will increase and it may not be possible to keep large blocks homogeneous.
iii) If the data on more than two plots is missing the statistical analysis becomes tedious and complicated.

Statistical analysis of R.B.D. for one observation per Experimental Unit

If in an R.B.D. a single observation is made on each of the experimental units, then it's analysis is same as ANOVA for fixed effect model for a two-way classified data with one observation per cell.

Null Hypothesis

H_{01}: the treatments are homogeneous i.e., $H_{01}: \tau_1 = \tau_2 = \tau_{3=\ldots=} \tau_t$

H_{02}: the blocks are homogeneous i.e., $H_{02}: b_1 = b_2 = b_{3=\ldots=} b_r$

Alternate hypothesis

H_{11}: At least two τ_i's are different

H_{12}: At least two b_j's are different

Least square estimates

The linear model is given by

$$y_{ij} = \mu + \tau_i + b_j + \epsilon_{ij}, \quad (i = 1,2\ldots t, j = 1,2\ldots r)$$

$$\epsilon_{ij} = y_{ij} - \mu - \tau_i - b_j$$

Applying the principle of least squares

$$E = \Sigma\Sigma\, \epsilon_{ij}^2 = \Sigma\, \Sigma\, (y_{ij} - \mu - \tau_i - b_j)^2$$

The normal equation for estimating the parameters μ, τ_i and b_j

$$\frac{\partial E}{\partial \mu} = 0 \rightarrow -2\, \Sigma\Sigma \left(y_{ij} - \mu - \tau_i - b_j\right) = 0 \quad \ldots\ldots (1)$$

$$\frac{\partial E}{\partial \tau_i} = 0 \rightarrow -2\, \Sigma\Sigma \left(y_{ij} - \mu - \tau_i - b_j\right) = 0 \quad \ldots\ldots (2)$$

$$\frac{\partial E}{\partial \mathrm{bj}} = 0 \rightarrow -2\, \Sigma\Sigma \left(y_{ij} - \mu - \tau_i - b_j\right) = \cdots \ldots (3)$$

Since $\Sigma\, \tau_i = 0 = \Sigma\, b_j$

We get from the above equations

$$\hat{\mu} = \frac{1}{\mathrm{tr}} \Sigma\Sigma y_{ij} = \bar{y}..$$

$$\widehat{\tau_i} = \frac{1}{\mathrm{t}} \Sigma y_{ij} - \hat{\mu} = \bar{y}_i. - \bar{y}..$$

$$\widehat{b_j} = \frac{1}{\mathrm{r}} \Sigma y_{ij} - \hat{\mu} = \bar{y}_{.j} - \bar{y}..$$

Thus, the linear $y_{ij} = \mu + \tau_i + b_j + \epsilon_{ij}$, model becomes

$$y_{ij} = \bar{y}.. + (\bar{y}_i. - \bar{y}..) + \left(\bar{y}_{.j} - \bar{y}..\right) + (y_{ij} - \bar{y}_i. - \bar{y}_{.j} + \bar{y}..)$$

Partitioning of the sum of squares

$$\Sigma\Sigma\,(y_{ij}-\bar{y}..)^2=\Sigma\Sigma[(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}..)+(\bar{y}_{i.}-\bar{y}..)+(\bar{y}_{.j}-\bar{y}..)]^2$$

$$\begin{aligned}&=\Sigma\Sigma(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}..)^2+\Sigma\Sigma\,(\bar{y}_{i.}-\bar{y}..)^2\\&+\Sigma\Sigma\,(\bar{y}_{i.}-\bar{y}..)^2\quad+2\Sigma\Sigma(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}..)(\bar{y}_{i.}-\bar{y}..)\\&+2\Sigma\Sigma(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}..)(\bar{y}_{.j}-\bar{y}..)\\&+2\Sigma\Sigma(\bar{y}_{i.}-\bar{y}..)(\bar{y}_{.j}-\bar{y}..)\end{aligned}$$

Now

$$\begin{aligned}\Sigma\Sigma(\bar{y}_{i.}-\bar{y}..)(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}..)&=\Sigma[(\bar{y}_{i.}-\bar{y}..)\Sigma(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}..)]\\&=\Sigma[(\bar{y}_{i.}-\bar{y}..)\{\Sigma(y_{ij}-\bar{y}_{i.})-\Sigma(\bar{y}_{.j}+\bar{y}..)\}]=0\end{aligned}$$

Since algebraic sum of deviations of a set of observations about their mean is 0. Similarly, all other product terms also vanish.

$$\Sigma\Sigma\,(y_{ij}-\bar{y}..)^2=\Sigma\,\Sigma(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}..)^2+b\Sigma\,(\bar{y}_{i.}-\bar{y}..)^2+t\Sigma\,(\bar{y}_{i.}-\bar{y}..)^2$$

Or $S_T^2 = S_t^2 + S_b^2 + S_e^2$

$$\Sigma\Sigma\,(y_{ij}-\bar{y}..)^2 \textit{ is the total sum of squares, } S_T^2$$

$$\Sigma\,(\bar{y}_{i.}-\bar{y}..)^2 \textit{ is the sum of squares due to treatments, } S_t^2$$

$$\Sigma\,(\bar{y}_{i.}-\bar{y}..)^2 \textit{ is the sum of squares due to blocks, } \quad S_b^2$$

$$\Sigma\,\Sigma(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}..)^2 \textit{ is the sum of squares due to error, } S_E^2$$

Degrees of freedom for various sum of squares

The total sum of squares S_T^2 being computed from rt quantities which are subject one linear constraint will carry (rt-1) d.f. similarly S_t^2 will have (t-1) d.f.,

S_r^2 will have (r-1) d.f., S_E^2 will have (r-1)(t-1) d.f.

Since N=rt

Thus the partitioning of d.f. is as follows (rt – 1) = (t – 1) + (r – 1) + (r -1)(t – 1)

Test Statistics

In order to obtain appropriate test statistics to test the hypothesis H_t and H_b we need the expectations of various mean sum of squares due to each of independent factors. Using the same notations for mean sum of squares as in the case of one-way classified data.

Mean Sum of Squares

M.S.S. due to treatment = $\frac{\mathbf{S_t^2}}{t-1} = s_t^2$

M.S.S. due to blocks = $\frac{\mathbf{S_b^2}}{r-1} = s_b^2$

Error Mean Sum Of Squares = $\frac{\mathbf{S_E^2}}{(t-1)(r-1)} = s_E^2$

ANOVA Table for R.B.D.

Sources of variation	d.f.	Sum of squares	Mean Sum of squares	Variance Ratio
Treatments	(t-1)	$\mathbf{S_T^2}$	$\frac{\mathbf{S_t^2}}{t-1} = s_t^2$	$F_t = \frac{S_t^2}{S_E^2}$
Blocks or replicates	(r – 1)	$\mathbf{S_b^2}$	$\frac{\mathbf{S_b^2}}{r-1} = s_b^2$	$F_b = \frac{S_b^2}{S_E^2}$
Error or residual	(t – 1) (r – 1)	$\mathbf{S_E^2}$	$\frac{\mathbf{S_E^2}}{(t-1)(r-1)} = s_E^2$	

Efficiency of RBD relative to CRD

The best method of calculating the efficiency of a design is to obtain the so called uniformity trial, i.e., the data obtained on a large number of experimental units with a common treatment.

Consider a design with t treatments, each replicated r times, Then ANOVA table for RBD is as given below

ANOVA table for RBD

Sources of variation	d.f.	Mean Sum of squares
Treatments	(t-1)	s_t^2
Blocks or replicates	(r – 1)	s_b^2
Error or residual	(t – 1) (r – 1))	s_E^2

Let the error sum of squares for RBD be s_E^2 and CRD be $s_E'^2$

If we apply uniformity trials to RBD which consists in using the same treatment on all the rt units, there are no treatment variations. Hence, consequently the treatment d.f. add to the error d.f..

Therefore error d.f. (due to treatment trials) = (t – 1) (r – 1) + (t – 1)

= (t - 1) (r – 1 + 1)

= r (t – 1)

Error sum of squares = r(t – 1) s_E^2

Now if we carry out the same experiment as a CRD, (on the same set of experimental units) then there is no variation between blocks. Hence the block d.f. and the block sum of squares add to the error d.f. and error sum of squares respectively.

Hence, for CRD

Error d.f. = r(t – 1) + (r – 1) = rt – 1

Therefore,

Error Mean Sum Of Squares for CRD is given by = $\boldsymbol{s_E'^2} = \frac{r(t-1)\mathbf{S_E^2}+(r-1)\mathbf{S_b^2}}{(rt-1)}$

Hence, the efficiency $E_1$of RBD relative to CRD is given by

$$E_1 = \frac{\mathbf{S}_{E\prime}^2}{\mathbf{S}_E^2} = = \frac{r(t-1)\mathbf{S}_E^2+(r-1)\mathbf{S}_b^2}{(rt-1)\mathbf{S}_E^2}$$

6·5·7. Estimation of Missing Value in R.B.D. Let the observation $y_{ij} = x$ (say) in the jth block and receiving the ith treatment be missing, as given in *Table 6·17.*

TABLE 6·17

		Treatments						Total
		1	2	...	i	...	t	
Blocks	1	y_{11}	y_{21}	...	y_{i1}	...	y_{t1}	$y_{\cdot 1}$
	2	y_{12}	y_{22}	...	y_{i2}	...	y_{t2}	$y_{\cdot 2}$
	⋮	⋮	⋮		⋮			⋮
	j	y_{1j}	y_{2j}	...	x	...	y_{ij}	$y_{\cdot j}' + x$
	⋮	⋮	⋮		⋮			⋮
	r	y_{1r}	y_{2r}	...	y_{ir}	...	y_{tr}	
	Total	$y_{1\cdot}$	$y_{2\cdot}$	...	$(y_{i\cdot}' + x)$	...	$y_{t\cdot}$	$y_{\cdot\cdot}' + x$

where $y_{i\cdot}'$ is total of *known* observations getting ith treatment,

$y_{\cdot j}'$ is total of *known* observation in jth block, and

$y_{\cdot\cdot}'$ is total of all the *known* observations

Total $S.S. = \sum\sum y_{ij}^2 - C.F. = x^2 + \text{constant}$, $w.r.t.\ x - C.F.$

$$S.S.T. = \frac{1}{r}\left[(y_{i\cdot}' + x)^2 + \text{constant } w.r.t\ \ x)\right] - C.F.$$

$$S.S.B. = \frac{1}{t}\left[(y_{\cdot j}' + x)^2 + \text{constant } w.r.t\ \ x\right] - C.F., \text{ where } C.F. = (y'_{\cdot\cdot} + x)^2/rt$$

$\therefore$ E = Residual S.S. = T.S.S. – S.S.B. – S.S.T.

$$= x^2 - \frac{1}{t}(y'_{\cdot j} + x)^2 - \frac{1}{r}(y'_{i\cdot} + x)^2 + \frac{(y'_{\cdot\cdot} + x)^2}{rt} + \text{Constant terms independent of } x.$$

We shall choose x such that E is minimum. For E to be minimum for variations in x, we must have

$$\frac{\partial E}{\partial x} = 0 = 2x - \frac{2}{t}(y'_{\cdot j} + x) - \frac{2}{r}(y'_{i\cdot} + x) + \frac{2}{rt}(y'_{\cdot\cdot} + x)$$

$$\Rightarrow \quad x\left(1 - \frac{1}{t} - \frac{1}{r} + \frac{1}{rt}\right) = \frac{1}{t}y'_{\cdot j} + \frac{1}{r}y'_{i\cdot} - \frac{1}{rt}y'_{\cdot\cdot}$$

$$\therefore \quad x = \frac{ry'_{\cdot j} + ty'_{i\cdot}' - y'}{(r-1)(t-1)} \qquad \ldots (6\text{-}3$$

TABLE 6·18