Statistical Analysis of RBD

*Statistical Analysis of RBD:

If in RBD, a single observation is made on each of the experimental units, then its analysis is analogous to ANOVA for fixed effect model for a two-way classified data with one observation per cell.

The model is

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

Where y_{ij} is the response or the yield of the experimental unit receiving the i^{th} treatment in the j^{th} block;

μ is the general mean effect

t_i is the effect due to the i^{th} treatment

b_j is the effect due to j^{th} block or replicate

$$e_{ij} \overset{i.i.d}{\sim} N(0, \sigma_e^2)$$

Where μ , t_i and b_j are constants so that $\sum_{i=1}^{t} t_i = 0 \quad and \quad \sum_{j=1}^{r} b_j = 0$

If we write $\sum_i \sum_j y_{ij} = G = Grand\ Total$

$\sum_j y_{ij} = T_i$ = Total for i^{th} treatment

$\sum_i y_{ij} = B_j$ = Total for j^{th} block

μ , t_i and b_j are estimated by the method of least squares

$$E = \sum_i \sum_j e_{ij}^2 = \sum_i \sum_j \left(y_{ij} - \mu - t_i - b_j\right)^2 \quad \ldots (1)$$

Differentiate with respect to μ

$$\frac{\partial E}{\partial \mu} = 0$$

$$2\sum_i \sum_j (y_{ij} - \mu - t_i - b_j)(-1) = 0$$

$$-2\sum_i \sum_j (y_{ij} - \mu - t_i - b_j) = 0$$

$$\sum_i \sum_j (y_{ij} - \mu - t_i - b_j) = \frac{0}{-2} = 0$$

$$\sum_i \sum_j y_{ij} - \sum_i \sum_j \mu - \sum_i \sum_j t_{i_i} - \sum_i \sum_j b_j) = 0$$

$$\sum_i \sum_j y_{ij} - tr\mu - r\sum_i t_i - t\sum_j b_j = 0$$

Where $\sum_i \sum_j y_{ij} = G$

$$G - tr\mu - r\sum_i t_i - t\sum_j b_j = 0 \quad \ldots (2)$$

Since $\sum_i t_i = 0$ and $\sum_j b_j = 0$ we get,

$$\hat{\mu} = \frac{G}{rt} = \overline{x}_{..}$$

Now, differentiating (1) w.r.to t_i we get,

$$\frac{\partial E}{\partial t_i} = 0$$

$$2\sum_j (y_{ij} - \mu - t_i - b_j)(-1) = 0$$

$$-2\sum_j (y_{ij} - \mu - t_i - b_j) = 0$$

$$\sum_j (y_{ij} - \mu - t_i - b_j) = \frac{0}{-2} = 0$$

$$\sum_j y_{ij} - \sum_j \mu - \sum_j t_i - \sum_j b_j) = 0$$

Where $\sum_j y_{ij} = T_i$

$$T_i = r\mu + rt_i - \sum_j b_j \quad \ldots(3)$$

$$\hat{t}_i = \frac{T_i}{r} - \hat{\mu}$$, Since $\sum_j b_j = 0$

$$= \bar{y}_{i.} - \bar{y}_{..} \quad \text{(where } \bar{y}_{i.} = \frac{T_i}{r} = \frac{\sum_j y_{ij}}{r} \text{ and } \bar{y}_{..} = \frac{G}{rt} = \frac{\sum_{i=1}^{t}\sum_{j=1}^{r} y_{ij}}{rt}\text{)}$$

Differentiating (1) w.r.to b_j ,

$$\frac{\partial E}{\partial b_j} = 0$$

$$2\sum_i (y_{ij} - \mu - t_i - b_j)(-1) = 0$$

$$-2\sum_i (y_{ij} - \mu - t_i - b_j) = 0$$

$$\sum_i (y_{ij} - \mu - t_i - b_j) = \frac{0}{-2} = 0$$

$$\sum_i y_{ij} - \sum_i \mu - \sum_i t_i - \sum_i b_j) = 0$$

$$\text{Where } \sum_i y_{ij} = B_j$$

$$B_j = t\mu + \sum_i t_i + tb_j \qquad \ldots.(4)$$

Hence,

$$\hat{b}_j = \frac{B_j}{t} - \hat{\mu}$$

$= \bar{y}_{.j} - \bar{y}_{..}$ (where $\bar{y}_{.j} = \frac{B_j}{t} = \frac{\sum_i y_{ij}}{t}$)

Now, partitioning total sum of squares in to various sum of squares as follows:

$$\sum_i^t \sum_j^r (y_{ij} - \bar{y}_{..})^2 = \sum_i^t \sum_j^r \left[(\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}) \right]^2$$

$$= r\sum_i (\bar{y}_{i.} - \bar{y}_{..})^2 + t\sum_j (\bar{y}_{.j} - \bar{y}_{..})^2 + \sum_i^t \sum_j^r (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$$

Since, sum of the deviation of the observation from its mean is always zero, the product terms shall vanish.

Total S.S = SSt + SSB + SSE

Where Total SS = $\sum_i^t \sum_j^r (y_{ij} - \bar{y}_{..})^2$ with (rt-1) d.f

SSt = Sum of squares due to treatments = $r\sum_i (\bar{y}_{i.} - \bar{y}_{..})^2$ with (t-1) d.f

SSB= Sum of squares due to Blocks = $t\sum_j (\bar{y}_{.j} - \bar{y}_{..})^2$ with (r-1) d.f

SSE = Sum of squares due to error = $\sum_i^t \sum_j^r (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$ with (r-1)(t-1) d.f.

Hence, the total sum of squares is partitioned or split into three sum of squares whose degrees of freedom add to the total degrees of freedom of Total ss.

i.e. (rt -1) = (t-1)+(r-1)+(r-1)(t-1)

*** ANOVA table for RBD:**

Source of variation	d.f	S.S.	M.S.S	F-ratio
Treatments	(t -1)	$S_t^2 = r\sum_i (\bar{y}_{i.} - \bar{y}_{..})^2$	$s_t^2 = S_t^2 / t-1$	$F_t = \frac{s_t^2}{s_E^2}$
Blocks	(r -1)	$S_b^2 = t\sum_j (\bar{y}_{.j} - \bar{y}_{..})^2$	$s_b^2 = S_b^2/r-1$	$F_b = \frac{s_b^2}{s_E^2}$
Error	(r-1)(t-1)	$S_E^2 = \sum_i^t \sum_j^r (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$	$s_E^2 = S_E^2 / (t-1)(r-1)$	
Total	(rt -1)			

Under the null hypothesis $H_{0t} = t_1 = t_2 = \ldots = t_i$ against the alternative that all t's are not equal the test statistic is :

$$F_t = \frac{s_t^2}{s_E^2} \sim F((t-1),(t-1)(r-1))$$

i.e., F_T follows F(central) distribution with [(t-1), (t-1)(r-1)] d.f. Thus if F_T is greater than tabulated F for [(t-1), (t-1)(r-1)] d.f, at certain level of significance, usually 5 % then we reject the null hypothesisH_{0t} and conclude that the treatments differ significantly. If F_t is less than tabulated value then F_T is not significant and we conclude that the data do not provide any evidence against the null hypothesis which may be accepted.

Similarly under the null hypothesis $H_{ob} = b_1 = b_2 = \ldots = b_r$, against the alternative that b' s are not equal, the test statistics is:

$$F_b = \frac{s_b^2}{s_E^2} \sim F((t-1),(t-1)(r-1))$$

Remark:

The following formulae reduce arithmetic to a great extent for the calculation of various ss:

1) Total S.S. = $\sum_{i=1}^{t}\sum_{j=1}^{r} y_{ij}^2 - \frac{G^2}{rt}$, where G = grand total

2) SSt = $\frac{1}{r}\sum_{i=1}^{t} T_{i.}^2 - C.F.$

3) SSB = $\frac{1}{t}\sum_{j=1}^{r} T_{.j}^2 - C.F$

4) SSE = Total SS- SSt –SSB

Where C.F. = $\frac{G^2}{rt}$

Note: If the mean effects for the treatments differ significantly, then we would be interested to find out which pair of means differ significantly. For this we compute the critical difference (C.D.) which is given by,

C.D.(treatment means) = $t_{(r-1)(t-1)}$ for $(\alpha/2)\%$* $\sqrt{\frac{2MSSE}{r}}$

Remark: It has been theoretically proved that,

1) $E\left(\frac{SSt}{t-1}\right) = r\sigma_t^2 + \sigma_e^2$

2) $E\left(\frac{SSb}{r-1}\right) = t\sigma_b^2 + \sigma_e^2$

3) $E\left(\frac{SSE}{(r-1)(t-1)}\right) = \sigma_e^2$

In this field there is a slope that causes a fertility gradient. The field is more fertile at the bottom than at the top. Different colors represent different treatments; each horizontal row represents a block. There are 4 blocks [I-IV] and 4 treatments [A (red); B (blue); C (yellow); D (green)] in this example.

Block	Treatment			
I	A	B	C	D
II	D	A	B	C
III	B	D	C	A
IV	C	A	B	D