Randomised Block Design.

Analysis of Variance for two Factors or Classification

Let the N variate values $\{x_{ij}\}$ (representing the yield of paddy) be classified according to two factors. Let there be 'h' rows (blocks) representing one factor of classification (soil fertility) and 'k' columns representing the other factor (treatment), so that $N = hk$.

We wish to test the null hypothesis that the rows and columns are homogeneous viz., there is no difference in the yields of paddy between the various rows and between the various columns.

Let x_{ij} be the variate value in the i^{th} row and j^{th} column.

Q = total variation.
Q_1 = sum of the squares due to the variations in the rows.
Q_2 = that in the columns and
Q_3 = that due to the residual variations.

Proceeding as in one factor of classification, we can prove that $\frac{Q_1}{h-1}, \frac{Q_2}{k-1}$, $\frac{Q_3}{(h-1)(k-1)}$ and $\frac{Q}{hk-1}$ are unbiased estimates of the population variance σ^2 with degrees of freedom $h-1, k-1, (h-1)(k-1)$ and $(hk-1)$ respectively. If the sampled population is assumed normal, all these estimates are independent.

$\therefore \frac{Q_1/(h-1)}{Q_3/(h-1)(k-1)}$ follows a F-distribution with $\{h-1, (h-1)(k-1)\}$ degrees of freedom and $\frac{Q_2/(k-1)}{Q_3/(h-1)(k-1)}$ follows a F-distribution with $\{k-1, (h-1)(k-1)\}$ degrees of freedom. Then the F-tests are applied as usual and the significance of difference between rows and between columns is analysed.

Table 10.2 The ANOVA table for the two factors of classifications

S.V.	*S.S.*	*d.f.*	*M.S.*	*F*
Between rows	Q_1	$h-1$	$Q_1/(h-1)$	$\left[\frac{Q_1/(h-1)}{Q_3/(h-1)(k-1)}\right]^{\pm 1}$
Between columns	Q_2	$k-1$	$Q_2/(k-1)$	$\left[\frac{Q_2/(k-1)}{Q_3/(h-1)(k-1)}\right]^{\pm 1}$
Residual	Q_3	$(h-1)(k-1)$	$Q_3/(h-1)(k-1)$	–
Total	Q	$hk-1$	–	–

Note *The following working formulas that can be easily derived may be used to compute Q, Q_1, Q_2 and Q_3:*

1. $Q = \sum\sum x_{ij}^2 - \frac{T^2}{N}$, *where* $T = \sum\sum x_{ij}$

2. $Q_1 = \frac{1}{k}\sum T_i^2 - \frac{T^2}{N}$, *where* $T_i = \sum_{j=1}^{k} x_{ij}$

3. $Q_2 = \frac{1}{h}\sum T_j^2 - \frac{T^2}{N}$, *where* $T_j = \sum_{i=1}^{h} x_{ij}$

4. $Q_3 = Q - Q_1 - Q_2$

It may be verified that $\sum_i T_i - \sum_j T_j = T.$

3. Latin Square Design (L.S.D.)

We consider an agricultural experiment, in which n^2 plots are taken and arranged in the form of an $n \times n$ square, such that the plots in each row will be homogeneous as far as possible with respect to one factor of classification, say, soil fertility and plots in each column will be homogeneous as far as possible with respect to another factor of classification, say, seed quality.

Then n treatments are given to these plots such that each treatment occurs only once in each row and only once in each column. The various possible arrangements obtained in this manner are known as Latin squares of order n. This design of experiment is called the Latin Square Design.

Analysis of Variance for Three Factors of Classifications

Let the $N\ (= n^2)$ variate values $\{x_{ij}\}$, representing the yield of paddy, be classified according to three factors. Let the rows, columns and letters stand for the three factors, say soil fertility, seed quality and treatment respectively.

We wish to test the null hypothesis that the rows, columns and letters are homogeneous. viz., there is no difference in the yield of paddy between the rows (due to soil fertility), between the columns (due to seed quality) and between the letters (due to the treatments).

Let x_{ij} be the variate value corresponding to the i^{th} row, j^{th} column and k^{th} latter.

Example 7

The following data represent the number of units of production per day turned out by 5 different workers using 4 different types of machines:

		Machine Type			
		A	*B*	*C*	*D*
Workers:	1	44	38	47	36
	2	46	40	52	43
	3	34	36	44	32
	4	43	38	46	33
	5	38	42	49	39

(a) Test whether the five men differ with respect to mean productivity.

(b) Test whether the mean productivity is the same for the four different machine types.

We subtract 40 from the given values and work out with new values of x_{ij}.

Worker	*A*	*B*	*C*	*D*	T_i	T_i^2/k	$\sum_j x_{ij}^2$
1	4	−2	7	−4	5	6.25	85
2	6	0	12	3	21	110.25	189
3	−6	−4	4	−8	−14	49.00	132
4	3	−2	6	−7	0	0	98
5	−2	2	9	−1	8	16.00	90
T_j	5	−6	38	−17	$T = 20$	$\sum \frac{T_j^2}{k} = 181.5$	594
T_j^2/h	5	7.2	288.8	57.8	$\sum T_j^2/h = 358.8$		
$\sum_i x_{ij}^2$	101	28	326	139	594		

$$Q = \sum\sum x_{ij}^2 - \frac{T^2}{N} = 594 - \frac{400}{20} = 574$$

$$Q_1 = \sum \frac{T_i^2}{k} - \frac{T^2}{N} = 181.5 - 20 = 161.5$$

$$Q_2 = \sum \frac{T_j^2}{h} - \frac{T^2}{N} = 358.8 - 20 = 338.8$$

$$Q_3 = Q - Q_1 - Q_2 = 574 - (161.5 + 338.8) = 73.7$$