MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI

DEPARTMENT OF STATISTICS

SUBJECT NAME : Design of Experiments

SUBJECT CODE : CST62

CLASS : III-B.Sc. STATISTICS

UNIT : IV

SYLLABUS

UNIT – IV

Missing plot technique – Meaning – Least square method of estimating missing Observations – one and two observations missing in RBD and LSD – Analysis of covariance technique in CRD and RBD(without derivation) – concept of Split-plot design

MISSING PLOT TECHNIQUES

In any design, the experiment in carried out by allocating various treatments over the experimental unit using the randomization principles and the yilds are noted for further analysis.

In some situation, the observations in one or two cells may be missing due to some reasons such as ignorance of the experimental carelessness of the field man.

In such circumstances, the missing values may be estimated using least square techniques, replacing the missing value, the analysis can be carried out in the usual way with the only differences that the total degrees of freedom is reduced by the no. of values being estimates.

Estimation of Missing Value in RBD

Let the observation $y_{ij} = x$ (say) in the j^{th} block and receiving the i^{th} treatment be missing.

Block	Treatment 1	2	...	i	...	t	Total
1	y_{11}	y_{12}	...	y_{i1}	...	y_{t1}	$y_{.1}$
2	y_{21}	y_{22}	...	y_{i2}	...	y_{t2}	$y_{.2}$
⋮	⋮	⋮		⋮		⋮	⋮
j	y_{1j}	y_{2j}	...	x	...	y_{tj}	$(y_{.j}'+x)$
⋮	⋮	⋮		⋮		⋮	⋮
r	y_{1r}	y_{2r}	...	y_{ir}	...	y_{tr}	$y_{.r}$
Total	$y_{1.}$	$y_{2.}$	...	$(y_{i.}'+x)$	...	$y_{t.}$	$y_{..}'+x$

where,

$y_{i.}'$ → Total of known observoration getting i^{th} treatment

$y_{.j}'$ → Total of known observation getting j^{th} block

$y_{..}'$ → Total of all the known observations.

Total Sum of Square

$$TSS = \sum\sum y_{ij}^2 - \text{Correction factor}$$

$$= x^2 + \text{constant with respect to } x - C.F.$$

Treatment Sum of square

$$SST = \frac{1}{r}\left[(y_{i.}' + x)^2 + \text{constant with respect to } x\right] - C.F.$$

Block sum of Square

$$SSB = \frac{1}{t}\left[(y_{.j}' + x)^2 + \text{constant with respect to } x\right] - C.F.$$

where,

$$\text{Correction factor} = \frac{(y_{..}' + x)^2}{rt}$$

Residual (Error) Sum of square

$$SSE = TSS - SST - SSB$$

$$= \left[x^2 + \text{constant w.r.to } x - C.F.\right] - \left\{\frac{1}{r}\left[(y_{i.}' + x^2)\right.\right.$$

$$\left.\left. + \text{constant w.r.to } x\right] - C.F.\right\} - \left\{\frac{1}{t}\left[(y_{.j}' + x)^2 + \text{cos.w.r.to } x \ldots\right.\right.$$

$$= \left[x^2 - CF\right] - \left[\frac{1}{r}(y_{i.}' + x)^2 - CF\right] - \left[\frac{1}{t}(y_{.j}' + x)^2 - CF\right]$$

$$+ \text{constant w.r.to } x \text{ (constant terms independent of } x\text{)}$$

$$= x^2 - \frac{1}{r}(y_{i.}' + x)^2 - \frac{1}{t}(y_{.j}' + x)^2 + \frac{(y_{..}' + x)^2}{rt}$$

we shall choose x, such that E is minimum

for E to be minimum for variations in x, so

we must have

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

$$2x - 2\frac{1}{r}(y_{i.}' + x) - 2\frac{1}{t}(y_{.j}' + x) + \frac{2(y_{..}' + x)}{rt} = 0$$

$$2x - \frac{2}{r} y_{i.}' - \frac{2}{r}x - \frac{2}{t} y_{.j}' - \frac{2}{t}x + \frac{2}{rt} y_{..}' + \frac{2x}{rt} = 0$$

$$2x - \frac{2x}{r} - \frac{2}{t}x + \frac{2x}{rt} = \frac{2}{r} y_{i.}' + \frac{2}{t} y_{.j}' - \frac{2}{rt} y_{..}'$$

$$\cancel{2}x\left(1 - \frac{1}{r} - \frac{1}{t} + \frac{1}{rt}\right) = \cancel{2}\left(\frac{y_{i.}'}{r} + \frac{y_{.j}'}{t} - \frac{y_{..}'}{rt}\right)$$

$$x\left(\frac{rt - t - r + 1}{\cancel{rt}}\right) = \left(\frac{t y_{i.}' + r y_{.j}' - y_{..}'}{\cancel{rt}}\right)$$

$$x(rt - t - r + 1) = t y_{i.}' + r y_{.j}' - y_{..}'$$

$$t(r-1) - r + 1$$

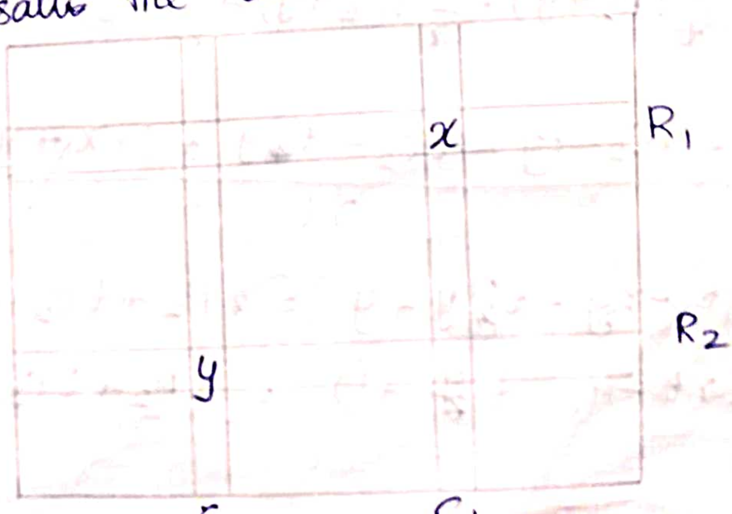
$$x[(r-1)(t-1)] = t y_{i.}' + r y_{.j}' - y_{..}'$$

$$x = \frac{t y_{i.}' + r y_{.j}' - y_{..}'}{(r-1)(t-1)}$$

30/01/24 **Two Missing Observation in RBD**

For two missing values (say) x and y. Let R_1 and R_2 be the totals of known observations in the row containing x and y respectively. And then C_1 and C_2 be the totals of known observations in the column containing x and y respectively.

Let S be the total of all the known observations as usally the error sum of squares becomes,



$$E = x^2 + y^2 - \frac{1}{t}\left[(R_1+x)^2 + (R_2+y)^2\right] - \frac{1}{r}\left[(C_1+x)^2 + (C_2+y)^2\right]$$
$$+ \frac{1}{rt}(S+x+y)^2 + (\text{terms independent of } x \text{ and } y)$$

For a minimum of E subject to variations in x and y, we must have

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

$$\Rightarrow 2x - \frac{1}{t}\left[2(R_1+x)\right] - \frac{1}{r}2(C_1+x) + \frac{1}{rt}2(S+x+y) = 0$$

$$\Rightarrow 2\left[x - \frac{1}{t}(R_1+x) - \frac{1}{r}(C_1+x) + \frac{1}{rt}(S+x+y)\right] = 0$$

$$x - \frac{1}{t}(R_1+x) - \frac{1}{r}(C_1+x) + \frac{1}{rt}(S+x+y) = 0$$

$$\frac{rtx - rR_1 - rx - tC_1 - tx + S + x + y}{rt} = 0$$

$$xrt - rx - tx + x = rR_1 + tC_1 - S - y$$

$$x(rt - r - t + 1) = rR_1 + tC_1 - S - y$$

$$x\{(r-1)(t-1)\} = rR_1 + tC_1 - S - y$$

$$\boxed{x = \frac{rR_1 + tC_1 - S - y}{(r-1)(t-1)}}$$

Solving $\frac{\partial E}{\partial y}$ similarly, we get

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

$$2y - \frac{1}{t}2(R_2+y) - \frac{1}{r}2(C_2+y) + \frac{1}{rt}2(S+x+y) = 0$$

$$2\left\{y - \frac{1}{t}(R_2+y) - \frac{1}{r}(C_2+y) + \frac{1}{rt}(S+x+y)\right\} = 0$$

$$\frac{yrt - rR_2 - ry - tC_2 - ty + S + x + y}{rt} = 0$$

$$yrt - ry - ty + y = rR_2 + tC_2 - S - x$$

$$y(rt - r - t + 1) = rR_2 + tC_2 - S - x$$

$$y[(r-1)(t-1)] = rR_2 + tC_2 - S - x$$

$$\boxed{y = \frac{rR_2 + tC_2 - S - x}{(r-1)(t-1)}}$$

NOTE:

Similarly, we can obtain the least square equations for the estimation of more than two missing observations.

Statistical Analysis

ANOVA is performed in the usual way after substituting the estimated values of the missing observations.

01/02/24 For each missing observation, one degrees of freedom is subtracted from total and consequently from error degrees of freedom.

The adjusted treatments sum of square is obtained by subtracting the adjustment ~~error~~ factor.

$$\frac{y'_{.j} + t y'_{i.} - y'_{..}}{t(t-1)(r-1)^2}$$

If the treatment show significant effect, then the standard error of the difference between two treatment means.

i) $\sqrt{\frac{2SE^2}{r}}$

If none of the treatments containing the missing value.

ii) $\left\{ SE^2 \left[\frac{2}{r} + \frac{t}{r(r-1)(t-1)} \right] \right\}^{1/2}$

If one of the treatments corresponds to the missing observation.

Estimating One missing observation in LSD

Let us suppose that in $m \times m$ Latin Square observation occuring in the i^{th} row, j^{th} column and receiving the k^{th} treatment is missing.

Let us assume that its value is x.

(ie) $y_{ijk} = x$

R = Total of known observation in the i^{th} row (ie) the row containing x.

C = Total of known observation in the j^{th} column (i.e) the column containing x.

T = Total of known observation receiving k^{th} treatment. (ie) Total of all known treatment values containing x.

S = Total of known observation, then we have

Total sum of square

$$TSS = x^2 + \text{constant w.r. to } x - \frac{(S+x)^2}{m^2}$$

Row sum of square

$$SSR = \frac{(R+x)^2}{m} + \text{constant w.r to } x - \frac{(S+x)^2}{m^2}$$

Column sum of square

$$SSC = \frac{(C+x)^2}{m} + \text{constant w.r. to } x - \frac{(S+x)^2}{m^2}$$

Treatment sum of square

$$SST = \frac{(T+x)^2}{m} + \text{constant w.r. to } x - \frac{(S+x)^2}{m^2}$$

Error sum of square

E = residual sum of square (SSE)

$= TSS - SSR - SSC - SST$

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

we shall choose x, such that E is minimum, so for E to be minimum for variation in x, then we must have $\frac{\partial E}{\partial x} = 0$

$$2x - \frac{1}{m}\left[2(R+x) + 2(C+x) + 2(T+x)\right] + 4\left(\frac{S+x}{m^2}\right) = 0$$

05/02/24

$$\Rightarrow 2\left[x - \frac{1}{m}(R+x+C+T+x) + \frac{2S}{m^2} + \frac{2x}{m^2}\right] = 0$$

$$\Rightarrow x - \frac{R}{m} - \frac{x}{m} - \frac{C}{m} - \frac{x}{m} - \frac{T}{m} - \frac{x}{m} + \frac{2S}{m^2} + \frac{2x}{m^2} = 0$$

$$\Rightarrow \frac{m^2x - Rm - Cm - Tm - 3xm + 2S + 2x}{m^2} = 0$$

$$\Rightarrow xm^2 - 3xm + 2x = Rm + Cm + Tm - 2S$$

$$x(m^2 - 3m + 2) = Rm + Cm + Tm - 2S$$

$$\boxed{x = \frac{Rm + Cm + Tm - 2S}{(m-1)(m-2)}}$$

After inserting the estimated value for missing observation, we perform the usual analysis of variance subtracting one degrees of freedom for total sum of square and consequently for error sum of square adjusted treatment sum of square is obtained by subtracting the quality.

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

From the treatment sum of square.

If the treatment show significant effect, then the standard error of the difference between two treatment mean,

i) $SE = \sqrt{\frac{2}{m}}$

If none of the treatment contains the missing value.

ii) $SE = \left[\frac{2}{m} + \frac{1}{(m-1)(m-2)}\right]^{1/2}$

If one of the treatments corresponds to the missing observation.

NOTE :

The same procedure will be followed for estimating more than 1, (i.e) k missing values and then missing values are obtained by solving k-equation simultaneously.

Analysis of Covariance

Analysis of covariances is used to test the main and interaction effects of categorical variables on a continuous dependent variable, controlling for the effects of selected other continuous variables, which covary with the dependent. The control variables are called the "covariates".

Analysis of Covariance (ANCOVA) is used in examining the differences in the mean values of the dependent variables that are related to the effect of the controlled independent variables while taking into account the influence of the uncontrolled independent variables.

Uses

To control experimental error and to adjust treatment means for the value of the covariate.

* To estimate missing data
* To aid in the interpretation of experimental results.

CRD

* Analysis of covariance (ancova) is a blend of regression and analysis of variance (anova) used for upgrade precision of an experiment.
* Ancova can be used for all experimental design including completely randomized design.
* The completely randomized design is the response are randomly intrusted to treatments.

Steps in Covariance Analysis (Randomised Complete Block Design)

1. Construct ANOVA tables as RCBD for X, independent variable or covariate, and for Y, dependent variable.
2. Check for treatment effect on X and on Y using F-test.
3. Calculate sums of cross-products
4. Construct Analysis of covariance table including sums of squares for X and Y, and sums of cross-products. Include Trt + Err df, SSX, SP and SSY.
5. Calculate SS Regr (adj for trt) and SS Dev Regr
6. Calculate SS Regr (trt + err) and SSTrt (adj for regr)
7. Complete the Analysis of covariance table and test MSRegr (adj for trt) and MSTr (adj for regr) against MSDevRegr (the remaining error).
8. Adjust treatment means

Analysis of Covariance for Two Way Classification (Random Block desisn) with one concomitant variable.

Suppose we want to compare v treatments, each treatment replicated r times so that total number of experimental units is $n = vr$. Suppose that the experiment is conducted with Randomized Block Design (RBD) layout.

Assuming a linear relationship between the response variable (y) and concomitant variable (x) the appropriate statistical model for ANOCOVA for RBD (with one concomitant variable) is:

$$y_{ij} = \mu + \alpha_i + \theta_j + \beta(x_{ij} - \bar{x}_{..}) + e_{ij} \quad (1)$$

where,

$\Rightarrow \mu$ is the general mean effect.

$\Rightarrow \alpha_i$ is the (fixed) additional effect due to the i^{th} treatment. $(i = 1, 2, \ldots v)$

$\Rightarrow \theta_j$ is the (fixed) additional effect due to the j^{th} block; $(j = 1, 2, \ldots r)$

$\Rightarrow \beta$ is the coefficient of regression of y on x

$\Rightarrow X_{ij}$ is the value of the concomitant variable corresponding to the response variable Y_{ij} and e_{ij} is the random error effect so that

$$\sum_{i=1}^{v} \alpha_i = 0, \quad \sum_{j=1}^{r} \theta_j = 0, \quad e_{ij} \overset{i.i.d}{\sim} N(0, \sigma_e^2)$$

Split Plot design

In field experiments certain factors may require larger plots than for others. For example, experiments on Irrigation, till age, etc., requires larger areas. On the other hand experiments on fertilizers, etc may not require larger areas. To accomodate factors which require different sizes of experimental plots in the same experiment, split plot design has been evolved.

In this design, larger plots are taken for the factor which requires larger plots. Next each of the larger plots is split into smaller plots to accomodate the other factor. The different treatments are allotted at random to their respective plots. Such arrangement is called Split Plot design.

In split plot design the larger plots are called main plots and smaller plots within the larger plot are called as sub-plots. The factor levels allotted to the main plots are main plot treatments and the factor levels allotted to sub plots are called as sub plot treatments.

Layout and Analysis of Variance table.

First the main plot treatment and sub plot treatment are usually decided based on the needed precision. The factor for which greater precision is required is assigned to the sub plots.

The replication is then divided into number of main plots equivalent to main plot treatments. Each main plot is divided into subplots depending on the number of sub plot treatments. The main plot treatments are allocated at random to the main plot as in the case of RBD. Within each main plot the subplot treatments are allocated at random as in the case of RBD. Thus, randomization is done in two stages. The same procedure is followed for all the replications independently.

The analysis of variance will have two parts, which correspond to the main plot and sub-plots. For the main plot analysis, replication × main plot treatment table is formed.

From this two-way table sum of squares for replication, main plot treatments and error (a) are computed. For the analysis of sub-plot treatments, main plot × sub-plot treatments table is formed. From this table the sums of squares for sub-plot treatments and interaction between main plot and sub-plot treatments are computed. Error (b) sum of squares is found out by residual method. The analysis of variance table for a split plot design A with 'm' main plot treatments and 's' sub-plot treatments with factor B is given below,

Source of variation	d.f	S.S	MS	F
Replication	$r-1$	RSS	RMS	$\frac{RMS}{EMS(a)}$
A	$m-1$	ASS	AMS	$\frac{AMS}{EMS(a)}$
Error (a)	$(r-1)(m-1)$	ESS (a)	EMS (a)	
B	$S-1$	BSS	BMS	$\frac{BMS}{EMS(b)}$
AB	$(m-1)(S-1)$	ABSS	ABMS	$\frac{ABMS}{EMS(b)}$
Error(b)	$m(r-1)(S-1)$	ESS(b)	EMS(b)	
Total	$rms-1$	TSS		

Treatment Combination	Replication R_1	R_2	R_3	....	Total
A_0B_0	a_0b_0	a_0b_0	a_0b_0	....	T_{00}
A_0B_1	a_0b_1	a_0b_1	a_0b_1	....	T_{01}
A_0B_2	a_0b_2	a_0b_2	a_0b_2	....	T_{02}
Sub Total	A_{01}	A_{02}	A_{03}	....	T_0
A_1B_0	a_1b_0	a_1b_0	a_1b_0	....	T_{10}
A_1B_1	a_1b_1	a_1b_1	a_1b_1	....	T_{11}
A_1B_2	a_1b_2	a_1b_2	a_1b_2	....	T_{12}
Sub Total	A_{11}	A_{12}	A_{13}	....	T_1
⋮	⋮	⋮	⋮	⋮	⋮
Total	R_1	R_2	R_3	....	GT

Complete $C.F = \frac{(GT)^2}{r \times m \times s}$

$TSS = [(a_0b_0)^2 + (a_0b_1)^2 + (a_0b_2)^2 + \dots] - CF$

Form A×R table and calculate RSS, ASS and Error(a) SS

Treatment	Replication				Total
	R_1	R_2	R_3	...	
A_0	A_{01}	A_{02}	A_{03}	...	T_0
A_1	A_{11}	A_{12}	A_{13}	...	T_1
A_2	A_{21}	A_{22}	A_{23}	...	T_2
⋮	⋮	⋮	⋮	⋮	⋮
Total	R_1	R_2	R_3	....	GT

$$RSS = \left(\frac{R_1^2 + R_2^2 + R_3^2 + \dots}{m.s}\right) - CF$$

$$ASS = \left(\frac{T_0^2 + T_1^2 + T_3^2 + \dots}{r.s}\right) - CF$$

$$A \times R \text{ table } SS = \left(\frac{A_{01}^2 + A_{02}^2 + A_{03}^2 + \dots}{b}\right) - CF.$$

Error (a) SS = AXRTSS − RASS − ASS.

From A×B Table and calculate BSS, A×B SS and Error (b) SS

Treatment	Replication				Total
	B_0	B_1	B_2	...	
A_0	T_{00}	T_{01}	T_{02}	...	T_0
A_1	T_{10}	T_{11}	T_{12}	...	T_1
A_2	T_{20}	T_{21}	T_{22}	...	T_2
⋮	⋮	⋮	⋮	⋮	⋮
Total	C_0	C_1	C_2		GT

$$BSS = \left(\frac{C_0^2 + C_1^2 + C_2^2 + \ldots}{r.m}\right) - CF$$

$$A \times B \text{ table } SS = \left(\frac{T_0^2 + T_1^2 + T_2^2 + \ldots}{r}\right) - CF$$

$$ABSS = A \times B \text{ Table } SS - ASS - ABSS$$

$$\text{Error (b) } SS = \text{Table } SS - ASS - BSS - ABSS - \text{Error (a) } SS$$

pdf

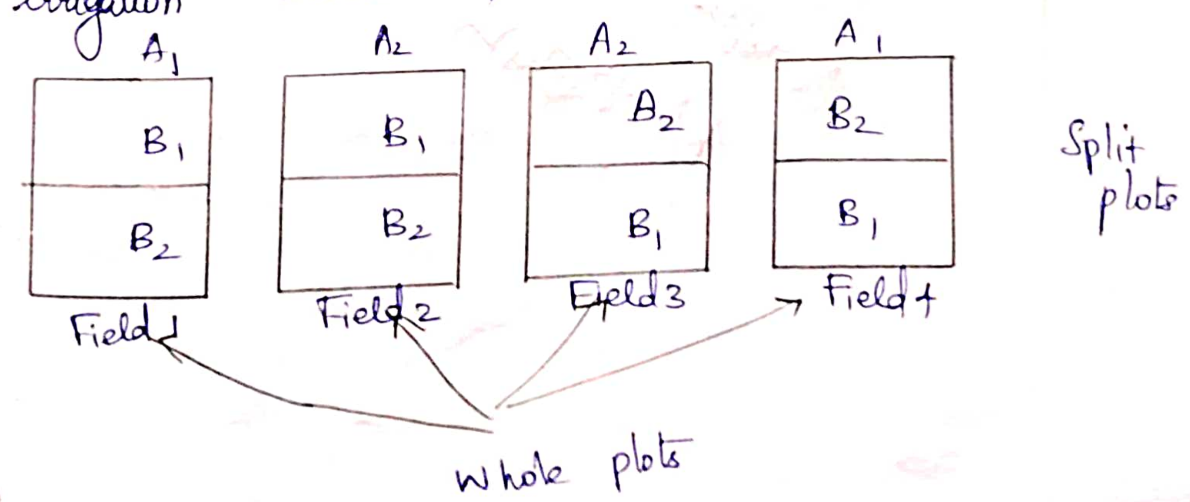
Then, complete the ANOVA table.

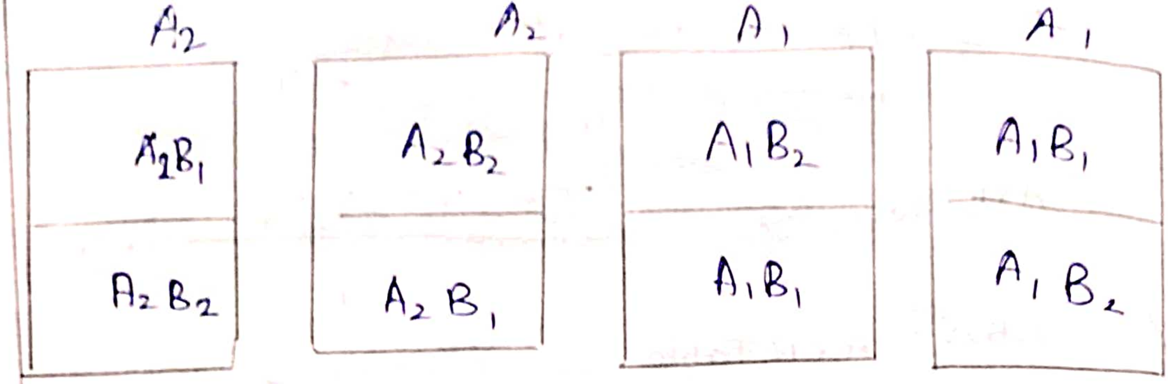
word

The split plot design is an experimental design that is used when a factorial treatment structure has two levels of experimental units. In this case of the split plot design, two levels of randomization are applied to assign experimental units to treatments.

The split plot design is an experimental design that is used when a factorial treatment structure ho

A typical example of a split plot design is an irrigation experiment where irrigation levels are applied to large areas, and factors like varieties and fertilizers are assigned to smaller areas within particular irrigation treatments.





4 Big plots = Whole plots

Irrigation (A) < I_1 (A_1), I_2 (A_2)

Why we use this,

It is used when some factors are harder (or more expensive) to vary than others. Basically, a split plot design consists of two experiments with different experimental units of different "size" experimental units, whereas other factors can be easily applied to "smaller" plots of land.

Verified by.
K.Raj 13/4/24

Text Books:

1. Das M.N and Giri N.C (1986) Design and Analysis of Experiments, Wiley Eastern, New Delhi.
2. Gupta, S.P. and Kapoor, V.K. (1978): Fundamentals of Applied Statistics, Sultan Chand & Sons.
3. Panneerselvam, R. (2012): Design and Analysis of Experiments, PHI.

Reference Books:

1. Rangaswamy, R (2014). Text book of Agricultural Statistics, New Age publishers
2. Montgomery, D (1972) Design of Experiments, John Wiley and Sons.
3. Kempthorne, (1956) Design and Analysis of Experiments, John Wiley. New York.