Design and Analysis of Experiments

Design of Single-Factor Experiments:

1- The Completely Randomized Design (CRD).
2- Randomized Complete Block Design (RCBD).
3- The Latin Square Design (LSD).

Completely Randomized Design (CRD):

1- CRD with one observation per experimental unit.
   a. In case of an equal number of replications.
   b. In case of unequal number of replications.
2- CRD with more than one observation per experimental unit.

-CRD with one observation per experimental unit and with equal number of replications of treatments:

The definition:

The design is used to compare treatments when the experimental units are essentially homogenous. Each treatment is applied at random to several experimental units.

Advantages:

1- CRD is the simplest design.
2- The statistical analysis is simple even in case the number of replicates are varies between one treatment and another.
3- The CRD is flexible in that number of treatments and replicates, if the numbers of homogenous experimental units are available.
4-The degrees of freedom for error are maximized. No other design with the same number of treatments and units provides greater error degree of freedom.
5- The loss of information due to missing data is small relative to losses with other designs.

Disadvantages:

1- The experimental units should be in a highly degree of homogenous.

2- The main disadvantage of this design is low precision if the experimental units receiving the same treatment are not uniform. In this case, real differences between treatments are difficult to detect.

Lay-out of experiment (Random distribution)

As the name states, all experimental units (treatment and replication combinations) are completely randomized within the whole experimental area. The steps in randomization are as follows:

1. Determine the total number of experimental units. For CRD with equal number of treatments per replication, it is the product of the number of treatments and the number of replicates. For CRD with unequal number of treatments, it is the sum of the treatments of all replicates.
2. Assign a plot number to each experimental unit consecutively.
3. Assign the plot numbers to the experimental plots using a randomization scheme. By using draw lots.
   (1) Prepare pieces of papers corresponding to the number of experimental units.
   (2) Write the plot number in each of the papers.
   (3) Mix the papers thoroughly in a container.
   (4) Without looking inside the container, draw a paper and assign the plot number on the first experimental unit. Suppose T5R1 was drawn first, it will be assigned to the first space in the experimental area.

T_5R_1					

(5) Without returning the paper previously drawn, continue drawing individual papers and assign the plot numbers until all the corresponding treatments have been assigned to all experimental units.

When the randomization is finished, the final lay-out may look like this:

T_5R_1	T_2R_6	T_4R_3	T_2R_5	T_3R_5	T_1R_1
T_2R_4	T_4R_4	T_5R_2	T_6R_5	T_5R_3	T_3R_6
T_6R_1	T_1R_4	T_1R_3	T_4R_2	T_6R_3	T_1R_2
T_3R_2	T_4R_5	T_6R_2	T_6R_4	T_6R_6	T_5R_5
T_1R_5	T_3R_3	T_5R_6	T_2R_3	T_5R_4	T_4R_1
T_3R_1	T_2R_1	T_4R_5	T_1R_5	T_3R_4	T_2R_2

The lay-out may also be like this:

T_1R_5	T_3R_3	T_5R_6	T_2R_3	T_5R_4	T_4R_1	T_2R_4	T_4R_4	T_5R_2	T_6R_5	T_5R_3	T_3R_4
T_5R_1	T_2R_6	T_4R_3	T_2R_5	T_3R_5	T_1R_1	T_6R_1	T_1R_4	T_1R_3	T_4R_2	T_6R_3	T_1R_2
T_3R_1	T_2R_1	T_4R_6	T_1R_6	T_3R_4	T_2R_2	T_3R_2	T_4R_5	T_6R_2	T_6R_4	T_6R_6	T_6R_5

or any other form provided the conditions are basically the same in the entire experimental unit.

Note that even under greenhouse condition, shading may differ from one site to another within the same greenhouse that is; plants placed near a wall may have less sunlight received compared to those place far from a wall. Likewise, plants placed near the window may receive more sunlight than those far from the window. Therefore, it will be better if all the treatments of an experiment are placed either near the window or away from the window whichever is suitable. It is not advisable that portion of the experiment is near the window while the other portion is far away from the window. The same principle applies to other conditions in the laboratory or greenhouse such as shading, temperature, wind direction, velocity, etc.

Steps of Analysis:

1- Symbolical representation of data:

Treatment (ti)	Observations (yij) r1	r2	r3	r4	Sum of treatments Yi.	Mean of treatments $\bar{y}i.$
t1	y11	y12	y13	y14	Y1.	$\bar{y}1.$
t2	y21	y22	y23	y24	Y2.	$\bar{y}2.$
t3	y31	y32	y33	y34	Y3.	$\bar{y}3.$
t4	y41	y42	y43	y44	Y4.	$\bar{y}4.$
					Y..= Grand Total	$\bar{y}$..= Grand mean

2- Linear model (Statistical model):

$y_{ij} = \mu + ti + e_{ij}$ [i=1,2,3.....t]
[j=1,2,3.....r]

Where,

yij= the value of the observation.

μ = effect of the overall mean.

$\mu = Y.. = \sum yij / tr = Y.. / tr$

$\hat{t}i$ = effect of the i^{th} treatment (i= t1, t2, t3.....tn)

$\hat{t}i = \bar{y}i. - \bar{y}..$

$\hat{e}ij$ = effect of the experimental error

$\hat{e}ij = yij - \bar{y}i.$

3- Test of Hypothesis:

1- Null hypothesis $H_o: \mu 1 = \mu 2 = ... = \mu t$

2- Alternative hypothesis $H_A: \mu 1 \neq \mu 2 \neq .... \neq \mu t$

4- Steps of ANOVA table (Analysis of Variance):

1- Determine the degree of freedom (d.f.):

a- Treatment d.f. = $t-1$

b- Error d.f. = $t(r-1)$

c- Total d.f. = $tr-1$

2- Determine the correction factor (C. F.)

C. F.= $(Y..)^2 / tr$

3- Estimation the sum of squares of each source of variation (SS):

a- Calculate the total sum of squares (SST):

$SST = \sum yij^2 - C.F.$

$= (y11^2 + y12^2 + \ldots\ldots\ldots + yij^2) - C.F.$

b- Calculate the treatments sum of squares (SSt):

$SSt = (\sum yi.^2 / r) - C.F.$

$= [(y1.^2 + y2.^2 + \ldots\ldots + yi.^2) / r] - C.F.$

c- Calculate the error sum of squares (SSe):

$SSe = SST - SSt$

4- Estimation the mean of squares of each variation source (MS):

a- Calculate the treatment mean of squares (MSt):

$MSt = SSt / t-1$

b- Calculate the error mean of squares (MSe):

$MSe = SSe / t(r-1)$

5- Determine the calculated F value (Cal. F.):

Cal. F.= MSt / MSe

6- Taking-out the Tabulated F value from F distribution Table (to know treatment d.f, error d.f and level of significant or probability level):

Alpha (α) = $P \leq 0.05$

Alpha (α) = $P \leq 0.01$

F- Value Table

Error d.f	Probability	Treatment d.f								
		1	2	3	4	5	6	7	8	9
1	0.100	39.86	49.50	53.59	...	...	...	...	...	...
	0.050	161.4	199.5	215.7	...	...	...	...	...	...
	0.025	647.8	799.5	864.2	...	...	...	...	...	...
	0.010	4052	4999.5	5403	...	...	...	...	...	...
	0.005	16211	20000	21615	...	...	...	...	...	...
2	0.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	0.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
	0.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
	0.010	98.5	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	0.005	198.5	199.0	199.2	199.2	199.3	199.3	199.4	199.4	199.4
3	0.100	5.54	5.46	5.39	5.34	5.31	5.29	5.27	5.25	5.24
.	0.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
.	0.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
.	0.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
.	0.005	55.55	49.80	47.47	46.16	45.39	44.84	44.43	44.13	43.88
.										
∞										

ANOVA Table

S.O.V.	d.f.	S.S.	M.S.	Cal.F.	Tab. F. 5%	1%	The result
Treats.	t-1	SSt	MSt	MSt/MSe	...	...	no significant
					*	...	significant
Error	t(r-1)	SSe	MSe		*	**	highly significant
Total	tr-1	SST					

5- The decision:

a- If the Cal. F. value less than Tab. F. value at both 5% and 1% probability level we say there are no significant differences (...) among treatments, hence we would be wise to accept the null hypothesis (H_o) and reject the alternative hypothesis (H_A).

b- If the Cal. F. greater than Tab. F. at 5% we say there are significant differences (*) among treatments, therefore we would be wise to accept the alternative hypothesis (H_A) and reject the null hypothesis (H_o).

c- If the Cal. F. greater than Tab. F. at probability level 1% we say there are a highly significant differences (**) among treatments, therefore we would be wise to accept the alternative hypothesis (H_A) and reject the null hypothesis (H_o).

6- Estimation of Variance components:

a- The variance of any observation in experiment:

$S^2_{yij} = MSe = \sigma^2_e$

b- The standard deviation of any observation:

$S_{yij} = \sqrt{Mse}$

c- The variance of any treatment mean:

$S^2_{\bar{y}i.} = MSe / r$

d- The standard deviation of any treatment mean:

$$S_{\bar{y}i.} = \sqrt{\frac{Mse}{r}}$$

e- The variance of any tow treatment mean:

$S^2(\bar{y}i. - \bar{y}i.) = 2MSe/r$

f- The standard deviation of any tow treatment mean:

$$S(\bar{y}i. - \bar{y}i.) = \sqrt{\frac{2Mse}{r}}$$

g- The coefficient of variation:

$C.V.\ \% = (\sqrt{Mse} / \bar{y}..) \times 100$

Exercise:

An experiment was carried out (in glass house) to compare four levels of nitrogen fertilizer (0, 2, 4 and 6 g/seedling) on radical length of wheat seedlings at probability 5%. A complete randomized design (CRD) was used. Each level being assigned to five pots (five replications). This table shows the length of radical after one month of treatments (cm/pot).

t1	18.4	t4	28.2	t3	24.1	t2	22.6	t2	23.2
t4	30.3	t1	16.1	t2	21.9	t4	27.4	t1	21.7
t3	25.2	t3	24.8	t4	26.4	t3	26.4	t4	34.8
t2	22.2	t1	17.3	t3	25.9	t2	23.9	t1	21.2

Treatments Ti	Observations (yij) r1	r2	r3	r4	r5	yi.	$\bar{y}i.$
t1	18.4	16.1	17.3	21.7	21.2	94.7	18.94
t2	22.2	21.9	22.6	23.9	23.2	113.8	22.76
t3	25.2	24.8	24.1	25.9	26.4	126.4	25.28
t4	30.3	28.2	26.4	27.4	34.8	147.1	29.42
						Y.. = 482	$\bar{y}..$ =24.1

$y_{ij} = \mu + ti + e_{ij}$ [i=1,2,3,4]

[j=1,2,3,4,5]

$H_o : \mu 1 = \mu 2 = \mu 3 = \mu 4$

$H_A : \mu 1 \neq \mu 2 \neq \mu 3 \neq \mu 4$

-Steps of analysis:

Treatment d.f = t -1 = 3

Error d.f = t(r -1) = 16

Total d.f = tr -1 = 19

C. F. = $(Y..)^2 /tr = (482)^2/ 20 = 11616.2$

SST=$\sum yij^2 - C.F. = (18.4^2 + 16.1^2 + ...34.8^2) - C. F. = 11980.8 - 11616.2 = 364.6$

SSt = $(\sum yi.^2 /r) - C.F. = [(94.7^2 + 113.8^2 + 126.4^2 + 147.1^2)/5] - C.F.$

$= 11906.78 - 11616.2 = 290.58$

SSe = SST –SSt = 364.6 – 290.58 = 74.02

MSt = SSt / t-1 = 96.86

MSe = SSe / t(r-1) = 4.626

Cal. F.= MSt / MSe = 20.938

Tab. F. = (3, 16, 5%) = 3.24

ANOVA Table

S.O.V.	d.f.	S.S.	M.S.	Cal. F.	Tab. F. 5%
Treats.	3	290.58	96.86	20.938*	3.24
Error	16	74.02	4.626		
Total	19	364.6			

The decision:

Because Cal. F. greater than Tab. F. at 5% we say there are significant differences (*) among treatments, therefore we would be wise to accept the alternative hypothesis (H_A) and refuse the null hypothesis (H_o).