

Chapter 2

Complete Randomized Design (CRD)



Complete Randomized Design (CRD)

Felix Kutsanedzie^{1*}; Sylvester Achio¹;

Edmund Ameko¹; Victoria Ofori²; Edith Mensah¹

¹Accra Polytechnic, GP 561, Accra, Ghana

²Agricultural Engineering Department, KNUST, Ghana

Abstract

Complete Randomized Design is one of the design analysis that is not well understood by researchers. However the need to understand such a design is very vital in the case where the experimental units being considered are homogeneous or uniform. This paper uses examples to explain the underlying principles and how the design is used for analysis.

Keywords

Randomization, Homogeneous, Treatment, Experimental Units, Factors, Levels, Replications

2.1 Introduction

Complete Randomized Design (CRD) is the simplest form of design used in experimental analysis. It involves randomization of treatments on homogeneous or uniform experimental units /plots uniform. Since in nature or reality, most experimental units are not homogeneous, it makes this design suited for large experiments vis-à-vis the homogeneity requirement of the experimental units. This presupposes that it is not suitable for the analysis of large field experiments. It is thus suited for small experiments.

In allocation or assignment of treatments to experimental units, randomization is used. Randomization involves the assigning of treatments to experiment units such that each treatment has equal chance of being assigned to units available. The randomization uses random tables and computer programmes to generate random numbers for the allocation of treatments to units. For instance, when one has five treatments – chloroquine (A), malarex (B), paracetamol (C), chamoquine (D) and panadol (E) are given to sterile rabbits(experimental units) to study their effects on them. Assuming that this experiment is to be replicated (repeated) five times, it means 5 (treatments) x 5 (replicates) totaling 25 (experimental units) from which data will be collected. The table below indicates how the randomization is done:

Table 2.1 Randomization of Five Treatments in a Complete Randomized Design (CRD).

A	B	C	D	E
B	A	D	C	D
C	D	E	A	C
D	E	B	B	A
E	C	A	E	B

Each of the boxes in the tables represent an experimental unit or plots (sterile rabbits) on which the various treatments (the medicines) denoted in the boxes are to be subjected to. This means that the researcher can table twenty (25) pieces of papers and label them with the letters A to E in turns until the twenty (25) pieces of papers are exhausted. The labeled pieces of papers are then folded, dropped in a cup and, shuffled. The pieces of papers are then picked randomly one after the order to denote the various plots or experimental units that the picked treatments would be applied to i.e. from 1st to the 25th experimental units. When this procedure is followed religiously, treatments are said to be randomized on the experimental units.

Once the experiment design is set, the treatment can then be allocated randomly to the experimental units according to the design. During this time onwards are the researcher is supposed to observe the observation units and to measure and determine responses of the experimental units to the treatments they have been subjected to.

2.2 Analysis of Data Obtained from the CRD

Assuming the treatments were subjected to the experiment units in order to determine their efficacies based on the time taken in days for a named symptom to be corrected, the researcher would have to observe the various treatments and record the time taken in days as summarized in the table below:

Data collected on time (measured in days) taken for a named symptoms on experimental units for each the respective replicates of each treatment considered in the CRD.

Table 2.2 Results Collected on responses of treatment to experimental units in CRD.

A = 3	B = 3	C = 3	D = 6	E = 3
B = 4	A = 4	D = 3	C = 4	D = 4
C = 5	D = 4	E = 2	A = 3	C = 4
D = 6	E = 3	B = 4	B = 3	A = 5
E = 3	C = 2	A = 5	E = 2	B = 3

In analyzing CRD designs after that had been collected from an experiment that have been carried, the must approach the Analysis of Variance (ANOVA) in terms of the following classification: designs having treatments with equal replication; designs having treatments with unequal replications, designs having treatments with equal number of samples per experimental unit; designs having number of subsamples that are unequal.

Design and Analysis of Variance (ANOVA) for treatments with equal number of replication

For the researcher to subject the data recorded in the Table 2.2, it must be re-summarized as below:

No/Replications	A	B	C	D	E
1	3	3	3	6	3
2	4	4	4	3	2
3	3	4	5	4	3
4	5	3	4	4	3
5	5	3	2	6	2

A = 3	B = 3	C = 3	D = 6	E = 3
B = 4	A = 4	D = 3	C = 4	D = 4
C = 5	D = 4	E = 2	A = 3	C = 4
D = 6	E = 3	B = 4	B = 3	A = 5
E = 3	C = 2	A = 5	E = 2	B = 3

The first step in the analysis is to put forth the hypothesis:

Before analyzing the data, there is the need to put forward the following hypothesis

$$H_0: \mu_A = \mu_B = \mu_C = \mu_D = \mu_E$$

$$H_A: \mu_A \neq \mu_B = \mu_C = \mu_D = \mu_E$$

H_0 : All the means of the treatments are equal

H_A : At least one of the means of the treatments is different

H_0 =Null hypothesis, H_A =Alternate hypothesis, μ = means of treatments

No/Replications	A	B	C	D	E
1	3	3	3	6	3
2	4	4	4	3	2
3	3	4	5	4	3
4	5	3	4	4	3
5	5	3	2	6	2
Tt =	20	17	18	23	13
$\Sigma Tt =$	91				
$CF = (\Sigma Tt)^2 / (r \times t)$	3.64				
$TSS = \Sigma((3)^2 + (3)^2 + \dots + (2)^2) - CF$	357.36				
$TtSS = \Sigma((20)^2 + (17)^2 + (13)^2) / r - CF$	338.56				
$ESS = TSS - TtSS$	18.8				

$$Tt_A = \sum (3 + 4 + \dots + 5) = 20$$

$$Tt_B = \sum (3 + 4 + \dots + 3) = 17$$

$$Tt_C = \sum (3 + 4 + \dots + 2) = 18$$

$$Tt_D = \sum(6 + 3 + \dots + 6) = 23$$

$$Tt_E = \sum(3 + 2 + \dots + 2) = 13$$

$$Tt_T = \sum(Tt_A + Tt_B + Tt_C + Tt_D + Tt_E) = 91$$

T_T=Total treatment, *ESS*=Error Sum of Squares, *TSS*= Total Sum of Squares, *TtSS*=Treatment Sum of Squares, *CF*= Correction Factor.

The second step is calculating of the Correction Factor:

$$CF = \frac{\sum(Tt_T)^2}{(r \times t)} = 331.24$$

The third step is calculating of the Total Sum of Squares:

$$TSS = \sum((3)^2 + (3)^2 + (3)^2 + (6)^2 + (3)^2 + (4)^2 \dots + (2)^2) - CF = 357.36$$

The fourth step is calculating of the Treatment Sum of Squares:

$$TtSS = \sum \frac{((20)^2 + (17)^2 + (18)^2 + (23)^2 + (13)^2)}{5} - CF = 338.56$$

The fifth step is calculating of the Error Sum of Squares:

$$ESS = TSS - TtSS = 357.36 - 338.56 = 18.8$$

T_T=Treatment Total, *TSS*=Total Sum of Squares, *TtSS*=Treatment Sum of Squares, *ESS*=Error Sum of Squares, *CF*=Correction Factor, *t*=number of treatment, *r* = number of replication.

The sixth step is completing the ANOVA table:

ANOVA TABLE					
Sources of Variation	df	SS	MS	Fcal.	Fcrit. (5%)
Treatment	$t-1 = 5 - 1 = 4$	338.56	84.64	90.04255	2.87
Error	$[(r \times t)-1-(t-1)] = 24 - 4 = 20$	18.8	0.94		
Total	$(r \times t) - 1 = (5 \times 5) - 1 = 24$	357.36			
Fcrit (5%) @ df of 4,20	= 2.87				
Fcrit (1%) @ df of 4,20	= 4.43				

df=degree of freedom, SS=Sum of Squares, MS=Mean Square.

The degree of freedom for the all items under the sources of variations is one minus the item: df of treatment is $5 - 1$; df for the total number of observation is $25-1$; the df of freedom for the error is $24 - 4$.

The Mean Square (MS) is the ratio of the SS to df: Treatment MS is calculated as:

$$\text{Treatment Mean Square (MS)} = \frac{338.56}{4} = 84.64$$

$$\text{Error Mean Square (MS)} = \frac{18.8}{20} = 0.94$$

The F calculated value is the ratio of Treatment Mean Square to the Error Mean Square:

$$F_{\text{cal.}} = \frac{\text{Treatment Mean Square}}{\text{Error Mean Square}} = \frac{84.64}{0.94} = 90.04255$$

The seventh step looking up the F-critical table to the F- critical values:

In looking at the F-critical table for the critical value, the level of significance or level of confidence is used. If the level of significance used is 1%, then the level of confidence is 99%; if the level of significance is 5%, the level of

confidence is 95%. Normally the 1% and 5% levels of significances are used depending on the precision of the work being undertaken.

Whether the 1% or 5% level of significance is used or not, the degree of freedom of the treatment variation is used against the degree of freedom of the error variation as (x, y), where the degree of freedom of the treatment is located horizontally (x-axis) against the degree of freedom of the error variation on the vertically (y-axis) and the intercept of these two gives the corresponding value for the F-critical or tabulated value.

Thus using the F-critical value from 5% (0.05) level of significance table for instance, one has to locate the degree of freedom of the treatment i.e. 4 horizontally on the 5% F- table, and the degree of freedom of the error i.e. 20 on the vertically. Where the two intercepts, the F- critical or tabulated value can be read as 2.87. When the same procedure is followed on the 1% F –table, at using the degree of freedom of the treatment as 4 against the degree of freedom of the error as 20, the F- critical or tabulated value of 4.43 is obtained. This is summarized as follows:

$$F_{\text{crit or tab. at 5\% (0.05)}} = 2.87$$

$$F_{\text{crit or tab. at 1\% (0.01)}} = 4.43$$

The eighth step is making the decision or conclusion:

The decision rule is summarized as follows:

If F calculated at the level of significance or level of confidence and degrees of freedoms stated is greater or equal to (\geq) the F critical or tabulated, the null hypothesis is rejected and therefore concluded that there is enough evidence from the data to support the rejection of the null hypothesis because the mean of the various treatments are significantly different. This is expressed mathematically as:

$$F_{cal} \geq F_{crit. \text{ or } tab. (5\%) \text{ at } 4,20}$$

$$F_{cal} \geq F_{crit. \text{ or } tab. (1\%) \text{ at } 4,20}$$

For instance in the case being considered:

$$F_{cal} (90.04) \geq 2.87, F_{crit. \text{ or } tab. (5\%) \text{ at } 4,20}$$

$$F_{cal} (90.04) \geq 4.43, F_{crit. \text{ or } tab. (1\%) \text{ at } 4,20}$$

Therefore the means of the various treatments being considered are significantly different at both 5% (0.05) and 1% (0.01)

However if the F calculated at the level of significance or level of confidence and degrees of freedoms stated is less than ($<$) the F critical or tabulated. We fail to reject the null hypothesis and conclude that there is enough evidence supporting the fact that the means of the various treatments are not significantly different. This is expressed as follows:

$$F_{cal} < F_{crit. \text{ or } tab. (5\%) \text{ at } 4,20}$$

$$F_{cal} < F_{crit. \text{ or } tab. (1\%) \text{ at } 4,20}$$

It should be noted whenever a level of significance is chosen; the p-value (probability value) is equal to the level of significance. Thus for 5%, it means the p-value = 0.05, at 1%, p-value = 0.01. This implies that when the F calculated \geq F critical, it is concluded that means are significantly different; hence the p-value will be less than the level of significance chosen i.e. < 0.05 and < 0.01 at 5% and 1% respectively.

Table 2.3 F critical table at 5% (0.05).

Critical values of F for the 0.05 significance					
	1	2	3	4	5
1	161.45	199.50	215.71	224.58	230.16
2	18.51	19.00	19.16	19.25	19.30
3	10.13	9.55	9.28	9.12	9.01
4	7.71	6.94	6.59	6.39	6.26
5	6.61	5.79	5.41	5.19	5.05
6	5.99	5.14	4.76	4.53	4.39
7	5.59	4.74	4.35	4.12	3.97
8	5.32	4.46	4.07	3.84	3.69
9	5.12	4.26	3.86	3.63	3.48
10	4.97	4.10	3.71	3.48	3.33
11	4.84	3.98	3.59	3.36	3.20
12	4.75	3.89	3.49	3.26	3.11
13	4.67	3.81	3.41	3.18	3.03
14	4.60	3.74	3.34	3.11	2.96
15	4.54	3.68	3.29	3.06	2.90
16	4.49	3.63	3.24	3.01	2.85
17	4.45	3.59	3.20	2.97	2.81
18	4.41	3.56	3.16	2.93	2.77
19	4.38	3.52	3.13	2.90	2.74
20	4.35	3.49	3.10	2.87	2.71
21	4.33	3.47	3.07	2.84	2.69

Table 2.4 F critical table at 1% (0.01).

Critical values of F for the 0.01 significance					
	1	2	3	4	5
1	4052.19	4999.52	5403.34	5624.62	5763.65
2	98.50	99.00	99.17	99.25	99.30
3	34.12	30.82	29.46	28.71	28.24
4	21.20	18.00	16.69	15.98	15.52
5	16.26	13.27	12.06	11.39	10.97
6	13.75	10.93	9.78	9.15	8.75
7	12.25	9.55	8.45	7.85	7.46
8	11.26	8.65	7.59	7.01	6.63
9	10.56	8.02	6.99	6.42	6.06
10	10.04	7.56	6.55	5.99	5.64
11	9.65	7.21	6.22	5.67	5.32
12	9.33	6.93	5.95	5.41	5.06
13	9.07	6.70	5.74	5.21	4.86
14	8.86	6.52	5.56	5.04	4.70
15	8.68	6.36	5.42	4.89	4.56
16	8.53	6.23	5.29	4.77	4.44
17	8.40	6.11	5.19	4.67	4.34
18	8.29	6.01	5.09	4.58	4.25
19	8.19	5.93	5.01	4.50	4.17
20	8.10	5.85	4.94	4.43	4.10
21	8.02	5.78	4.87	4.37	4.04

The ninth step is calculating the Coefficient of Variation:

$$\%CV = \frac{S}{\bar{X}} \times 100$$

$$\text{Error MS} = S^2$$

$$S = \sqrt{\text{Error MS}}$$

$$\%CV = \frac{\sqrt{\text{Error MS}}}{\bar{X}}$$

$$\%CV = \frac{\sqrt{0.94}}{3.64} \times 100 = 26.64$$

$$CV = 26.64\%$$

CV = coefficient of variation, S = Standard deviation

Error MS = Error Mean Square, \bar{x} = mean of treatment

The Coefficient of Variation is the ratio of the standard deviation to the mean of the treatments expressed in percentage. It is a measure of the consistency of the mean of the treatment or the variations in the mean of the treatment. It

implies that when the standard deviation increases the variations in the means of the treatment increases. The CV can also be taken as the measure of variations in the means of the treatments. Thus the lower the CV value, the more consistent or uniform the means of the treatment.

Design and Analysis of Variance (ANOVA) for treatments with unequal number of replication

No/Replications	A	B	C	D	E
1	3	3	3	6	3
2	4	4	4	3	2
3	3	4	5	4	3
4		3		4	3
5		3			2

This analysis is done when dealing with unequal number of replications of the treatment. In cases like this, the same preambles are used but with variations in the various formulae used in the computation. So we proceed to follow the same steps used earlier.

The first step in the analysis is to put forth the hypothesis:

Before analyzing the data, there is the need to put forward the following hypothesis

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_D = \mu_E$$

$H_1: \mu_1 \neq \mu_2 = \mu_3 = \mu_D = \mu_E$ (At least one of the treatment means differ from the others).

The second step is calculating of the Correction Factor:

No/Replications	A	B	C	D	E
1	3	3	3	6	3
2	4	4	4	3	2
3	3	4	5	4	3
4		3		4	3
5		3			2
Tt	10	17	12	17	13
$\sum Tt$	69				
$CF = (\sum Tt)^2 / (r)$	238.05				
TSS	16.95				
TtSS	7.13				
ESS	9.82				
Tt mean	3.45				

$$CF = \frac{\sum(Tt_T)^2}{r}$$

$$CF = \frac{(69)^2}{20} = 238.05$$

where $r = (r_A + r_B + r_C + r_D + r_E)$, $r =$ total number of observations for the treatments.

The third step is calculating of the Total Sum of Squares:

$$TSS = \sum ((3)^2 + (3)^2 + (3)^2 + (6)^2 + (3)^2 + (4)^2 \dots + (2)^2) - 238.05 = 16.95$$

The fourth step is calculating of the Treatment Sum of Squares:

$$TtSS = \sum \left(\frac{(Tt_A)^2}{r_A} + \frac{(Tt_B)^2}{r_B} + \frac{(Tt_C)^2}{r_C} + \frac{(Tt_D)^2}{r_D} + \frac{(Tt_E)^2}{r_E} \right) - CF$$

$$TtSS = \sum \left(\frac{(10)^2}{3} + \frac{(17)^2}{5} + \frac{(12)^2}{3} + \frac{(17)^2}{4} + \frac{(13)^2}{5} \right) - 238.05 = 7.13$$

$$TtSS = \sum 245.18 - 238.05 = 7.13$$

The fifth step is calculating of the Error Sum of Squares:

$$ESS = TSS - TtSS = 16.95 - 7.13 = 9.82$$

The sixth step is completing the ANOVA table:

ANOVA TABLE					
Sources of Variation	df	SS	MS	Fcal.	Fcrit. (5%)
Treatment	$t-1 = 5 - 1 = 4$	7.13	1.78	2.72	3.06
Error	$[(r-1)-(t-1)] = 19-4 = 15$	9.82	0.65		
Total	$r-1 = (20) - 1 = 19$	16.95			

The seventh step looking up the F-critical Table to the F-critical values:

This follows the earlier procedures used in looking up the critical values from the F critical Table

$$F_{crit. \text{ or tab. (5\%) at df of 4,15} = 3.06$$

$$F_{crit. \text{ or tab. (1\%) at df of 4,15} = 4.89$$

Critical values of F for the 0.05 significanc						Critical values of F for the 0.01 significanc					
	1	2	3	4	5		1	2	3	4	5
1	161.45	199.50	215.71	224.58	230.16	1	4052.19	4999.52	5403.34	5624.62	5763.65
2	18.51	19.00	19.16	19.25	19.30	2	98.50	99.00	99.17	99.25	99.30
3	10.13	9.55	9.28	9.12	9.01	3	34.12	30.82	29.46	28.71	28.24
4	7.71	6.94	6.59	6.39	6.26	4	21.20	18.00	16.69	15.98	15.52
5	6.61	5.79	5.41	5.19	5.05	5	16.26	13.27	12.06	11.39	10.97
6	5.99	5.14	4.76	4.53	4.39	6	13.75	10.93	9.78	9.15	8.75
7	5.59	4.74	4.35	4.12	3.97	7	12.25	9.55	8.45	7.85	7.46
8	5.32	4.46	4.07	3.84	3.69	8	11.26	8.65	7.59	7.01	6.63
9	5.12	4.26	3.86	3.63	3.48	9	10.56	8.02	6.99	6.42	6.06
10	4.97	4.10	3.71	3.48	3.33	10	10.04	7.56	6.55	5.99	5.64
11	4.84	3.98	3.59	3.36	3.20	11	9.65	7.21	6.22	5.67	5.32
12	4.75	3.89	3.49	3.26	3.11	12	9.33	6.93	5.95	5.41	5.06
13	4.67	3.81	3.41	3.18	3.03	13	9.07	6.70	5.74	5.21	4.86
14	4.60	3.74	3.34	3.11	2.96	14	8.86	6.52	5.56	5.04	4.70
15	4.54	3.68	3.29	3.06	2.90	15	8.68	6.36	5.42	4.89	4.56
16	4.49	3.63	3.24	3.01	2.85	16	8.53	6.23	5.29	4.77	4.44

The eighth step is making the decision or conclusion:

$$F_{cal. (2.72)} < (3.06), F_{crit. \text{ or tab. (5\%) at df of 4,15}$$

$$F_{cal.} (2.72) < (4.89), F_{crit. \text{ or tab.}} (1\%) \text{ at df of } 4,15$$

We therefore fail to reject the null hypothesis at both 5% (0.05) and 1% (0.01) level of significance and conclude that the means are not significantly different based on the data available.

The ninth step is calculating the Coefficient of Variation:

$$\%CV = \frac{\sqrt{\text{Error MS}}}{\bar{X}}$$

$$\%CV = \frac{\sqrt{0.65}}{3.45} \times 100 = 23.45$$

$$CV = 23.45\%$$

Design and Analysis of Variance (ANOVA) with equal number of samples per experimental units with treatments of equal number of replication

Consider an experimental design where three different concentrations (10M, 20M and 30M) of acetone are applied to dissolve a Perspex material chopped into five different thicknesses and replicated four times alongside taking the records of the time for the dissolution of chopped to be realized in minutes. This is an example of CRD design that requires ANOVA with equal number of samples per experimental unit with treatments of equal number of replication. The table below represents the summary of this information:

Conc. (M)	10M				20M				30M			
	Perspex"				Perspex"				Perspex"			
Thickness (inches)	1	2	3	4	1	2	3	4	1	2	3	4
1	4	5	6	4	5	6	7	8	9	4	5	5
2	5	4	6	4	3	6	7	7	4	6	5	5
3	6	7	7	5	6	6	7	5	5	5	4	7
4	4	7	8	6	4	5	5	6	6	6	4	6
5	7	8	9	4	5	6	7	4	7	7	8	9

The first step in the analysis is to put forth the hypothesis:

Before analyzing the data, there is the need to put forward the following hypothesis

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$H_1: \mu_1 \neq \mu_2 = \mu_3$ (At least one of the treatment means differ from the others)

Conc. (M)	10M				20M				30M			
	Perspex''				Perspex''				Perspex''			
Thickness (inches)	1	2	3	4	1	2	3	4	1	2	3	4
1	4	5	6	4	5	6	7	8	9	4	5	5
2	5	4	6	4	3	6	7	7	4	6	5	5
3	6	7	7	5	6	6	7	5	5	5	4	7
4	4	7	8	6	4	5	5	6	6	6	4	6
5	7	8	9	4	5	6	7	4	7	7	8	9
EU	26	31	36	23	23	29	33	30	31	28	26	32
T _t	116				115				117			
CF	2018											
Tt mean	5.8											
TSS	123.6											
TtSS	0.1											
EUSS	34.8											
EESS	34.7											
SESS	88.8											

The second step is calculating of the Correction Factor:

$$CF = \frac{\sum(Tt_T)^2}{rts}$$

$$CF = \frac{(348)^2}{(4)(3)(5)} = 2018.40$$

$r =$ number of replication, $t =$ number of treatment, $s =$ number of samples.

The third step is calculating of the Total Sum of Squares:

$$TSS = \sum ((4)^2 + (5)^2 + (6)^2 + (4)^2 + (5)^2 + (6)^2 \dots + (9)^2) - 2018.40 = 123.60$$

The fourth step is calculating of the Treatment Sum of Squares:

$$TtSS = \sum \left(\frac{(Tt_{t1})^2}{rs} + \frac{(Tt_{t2})^2}{rs} + \frac{(Tt_{t3})^2}{rs} \right) - CF$$

$$TtSS = \sum \left(\frac{(116)^2}{(4)(5)} + \frac{(115)^2}{(4)(5)} + \frac{(117)^2}{(4)(5)} \right) - 2018.40$$

$$TtSS = \sum 2018.5 - 2018.4 = 0.1$$

The fifth step is calculating the Sum of Squares among the Experimental Units:

$$EUSS = \sum \left(\frac{(EU_1)^2}{s} + \frac{(EU_2)^2}{s} + \frac{(EU_3)^2}{s} + \dots + \frac{(EU_{12})^2}{s} \right) - CF$$

$$EUSS = \sum \left(\frac{(26)^2}{5} + \frac{(23)^2}{5} + \frac{(31)^2}{5} \right) - 2018.40$$

$$EUSS = 2053.2 - 2018.40 = 34.8$$

The sixth step is calculating the Experimental Error Sum of Squares:

$$\text{Experimental Error SS (EESS)} = EUSS - TtSS$$

$$\text{Experimental Error SS (EESS)} = 34.8 - 0.1 = 34.70$$

The seventh step is calculating the Sampling Error Sum of Squares:

$$\text{Sampling Error SS (SESS)} = TSS - EUSS$$

$$\text{Sampling Error SS (SESS)} = 123.60 - 34.80 = 88.8$$

The eighth step is completing the ANOVA table:

ANOVA TABLE					
Sources of Variations	df	SS	MS	Fcal	Fcrit.
Treatment	$t - 1 = 3 - 1 = 2$	0.1	0.05	0.013	4.26(5%)
Experimental error	$(tr-1) - (t - 1) = (3 \times 4 - 1) - 2 = 9$	34.7	3.86		8.02(1%)
Sampling error	$(trs - 1) - (tr - 1) = (3 \times 4 \times 5 - 1) - 9 = 50$	88.8			
Total	$trs - 1 = 60 - 1 = 59$	123.6			

The seventh step looking up the F-critical Table to the F-critical values:

This follows the earlier procedures used in looking up the critical values from the F critical Table

$$F_{\text{crit. or tab. (5\%) at df of 2,9}} = 4.26$$

$$F_{\text{crit. or tab. (1\%) at df of 2,9}} = 8.02$$

The eighth step is making the decision or conclusion:

$$F_{\text{cal. (0.013)}} < (4.26), F_{\text{crit. or tab. (5\%) at df of 2,9}}$$

$$F_{\text{cal. (0.013)}} < (8.02), F_{\text{crit. or tab. (1\%) at df of 2,9}}$$

We therefore fail to reject the null hypothesis at both 5% (0.05) and 1% (0.01) level of significance and conclude that the means are not significantly different based on the data available.

The ninth step is calculating the Coefficient of Variation:

$$\%CV = \frac{\sqrt{\text{Error MS}}}{\bar{X}}$$

$$\%CV = \frac{\sqrt{3.86}}{5.8} \times 100$$

$$\%CV = \frac{1.96}{5.8} \times 100 = 33.87$$

$$CV = 33.87\%$$

Design and Analysis of Variance (ANOVA) with unequal number of subsamples

Taking for instance a design similar to the previously treated one where this time around the number of subsamples are unequal as indicated in the table, the following approach is used in analyzing it.

Conc. (M)	10M				20M				30M			
	Perspex''				Perspex''				Perspex''			
Thickness (inches)	1	2	3	4	1	2	3	4	1	2	3	4
1	4	5	6	4	5	6	7	8	9	4	5	5
2	5	4	6	4	3	6	7	7	4	6	5	5
3	6	3	7	5	6	7	7	5	5	4	4	7
4					4	5	5	6	4	4	4	6
5									6	5	8	5
TEU	15	12	19	13	18	24	26	26	28	23	26	28
T _T			59			94				105		
CF	1387											
Tt mean	5.37											
TSS	85											
TtSS	7											
EUSS	30											
EESS	23											
SESS	55											

The first step in the analysis is to put forth the hypothesis:

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$H_1: \mu_1 \neq \mu_2 = \mu_3$ (At least one of the treatment means differ from the others)

The second step is calculating of the Correction Factor:

$$CF = \frac{(\text{sum of observations})^2}{\text{number of observations}}$$

$$CF = \frac{(258)^2}{(48)}$$

$$CF = \frac{66564}{48} = 1387$$

The third step is calculating of the Total Sum of Squares:

$$TSS = \sum ((4)^2 + (5)^2 + (6)^2 + (4)^2 + (5)^2 + (6)^2 \dots + (5)^2) - 1387$$

$$TSS = 1472 - 1387 = 85$$

The fourth step is calculating of the Treatment Sum of Squares:

$$TtSS = \sum \left(\frac{(Tt_{t1})^2}{rs} + \frac{(Tt_{t2})^2}{rs} + \frac{(Tt_{t3})^2}{rs} \right) - CF$$

$$TtSS = \sum \left(\frac{(59)^2}{(4)(3)} + \frac{(94)^2}{(4)(4)} + \frac{(105)^2}{(4)(5)} \right) - 1387$$

$$TtSS = \sum 1394 - 1387 = 7$$

The fifth step is calculating the Sum of Squares among the Experimental Units:

$$EUSS = \sum \left(\frac{(EU_1)^2}{s_1} + \frac{(EU_2)^2}{s_2} + \frac{(EU_3)^2}{s_3} + \dots + \frac{(EU_{12})^2}{s_i} \right) - CF$$

$$EUSS = \sum \left(\frac{(15)^2}{3} + \frac{(18)^2}{4} + \frac{(28)^2}{5} \right) - 1387$$

$$EUSS = 1417 - 1387 = 30$$

EUSS = Sum of Squares Among Experimental Units

TEU = Total treatment on Experimental Units

The sixth step is calculating the Experimental Error Sum of Squares:

$$\text{Experimental Error SS (EESS)} = EUSS - TtSS$$

$$\text{Experimental Error SS (EESS)} = 30 - 7 = 23$$

The seventh step is calculating the Sampling Error Sum of Squares:

$$\text{Sampling Error SS (SESS)} = \text{TSS} - \text{EUSS}$$

$$\text{SamplingError SS(SESS)} = 85 - 30 = 55$$

ANOVA TABLE						
Sources of Variations	df	SS	MS	Fcal	Fcrit. (5%)	Fcrit (1%)
Treatment	$t - 1 = 3 - 1 = 2$	7	3.5	1.3696	4.26	8.02
Experimental error	$df(\text{EU}) - df(t) = (11 - 2) = 9$	23	2.56	2.5556		
Sampling error	$df(T) - df(\text{EU}) = (47 - 11) = 36$	55				
Total	$\text{Total observ.} - 1 = (48 - 1) = 47$	85				

df(EU)= degree of freedom of experimental unit.

df(T)= degree of freedom of total observation.

The seventh step looking up the F-critical Table to the F-critical values:

This follows the earlier procedures used in looking up the critical values from the F critical Table

$$\text{Fcrit. or tab. (5\%)} \text{ at } df \text{ of } 2,9 = 4.26$$

$$\text{Fcrit. or tab. (1\%)} \text{ at } df \text{ of } 2,9 = 8.02$$

The eighth step is making the decision or conclusion:

$$\text{Fcal. (1.37)} < (4.26), \text{ Fcrit. or tab. (5\%)} \text{ at } df \text{ of } 2,9$$

$$\text{Fcal. (1.37)} < (8.02), \text{ Fcrit. or tab. (1\%)} \text{ at } df \text{ of } 2,9$$

We therefore fail to reject the null hypothesis at both 5% (0.05) and 1% (0.01) level of significance and conclude that the means are not significantly different based on the data available.

Bibliography

- [1] Brownlee, K. A. (1960). *Statistical theory and methodology in science and engineering*. New York: Wiley.
- [2] Campbell, D., Stanley J. (1963). *Experimental and quasi-experimental designs for research and teaching*. In Gage (Ed.), *Handbook on research on teaching*. Chicago: Rand McNally & Co.
- [3] Winer, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.
- [4] Cochran, W. G., G. M. (1957). *Experimental Designs*, John Willey & Sons, N.Y.
- [5] Le Clerg, E. L., Leonard, W. H., Clark A. G. (1966). *Field Plot Technique*. Burgess Pub. Co., Minn., USA.

