Chapter 9

EXPERIMENTAL DESIGN

Definition of Terms

Experiment: a scientific investigation conducted under controlled conditions where *treatments* are applied on the experimental unit and their *effects* are measured and compared

Experimental unit: denotes the group of material to which a treatment is applied in a single trial of the experiment (ex: plot of land, patient in a hospital, group of pigs in a pen, batch of seeds)

Experimental error: extraneous variations in the values of the response variable which tend to mask the true effects of the treatments due to the following: (i) the inherent variations among experimental units in terms of the values of the extraneous variables and (ii) lack of uniformity in the physical conduct of the experiment resulting to variations in the values of more extraneous variables

Experimental design: a plan for conducting the experiment that includes the following: (i) selection of treatments; (ii) choice of experimental units; (iii) rules for applying the treatments; (iv) measurement of responses; (v) statistical analysis of data.

Basic Principles of Experimental Design

- ***Replication***

  This is the application of a treatment to more than one experimental unit. Replication allows us to estimate the size of the experimental error so that we can identify the sources of variation in the values of the response variable (Treatment vs experimental error). Moreover, increasing the number of replications is one factor that we consider in reducing the experimental errors.
- ***Blocking***

  This is the allocation of the experimental units into more or less homogeneous clusters called 'blocks'. The objective of blocking is to have units within a block as uniform as possible. This is so that units to which one treatment is applied are closely comparable with those units in the same block to which other treatments were applied.
- ***Randomization***

  This is a process by which the allocation of the experimental units and the order in which the individual runs or trials of the experiment are to be performed are determined using a randomization procedure. Through proper randomization, the effects of extraneous factors that were not consciously controlled are "averaged out"; that is, they cancel each other. Randomization also assures the independence of the observations.

Completely Randomized Design (CRD)

- *Description:*

  If there are r treatments then these treatments are allotted to the n units in such a way that if there are n_i replicates for the i^{th} treatment then all possible ways of grouping the n units into r groups of $n_1, n_2, \ldots n_r$ units must be given equal chances of selection.

- *Randomization:*

  Using a randomization mechanism, assign a distinct random number to each experimental unit. The n_1 units that were assigned the smallest random numbers will be given Treatment 1; the n_2 experimental units that were assigned the next smallest random numbers will be given Treatment 2, and so on.

Example 1

The effective life of a cutting tool installed in a machine is thought to be affected by the cutting speed. Three cutting speeds (125 in/min, 150 in/min, and 175 in/min) were considered in an experiment with 3 replicates each.

Response variable:	Length of life
Factor (explanatory variable):	Cutting speed
3 Treatments:	(1) 125 in/min
	(2) 150 in/min
	(3) 175 in/min
Experimental units:	cutting tools

Example 1 (cont'd)

- Using CRD with 3 replicates, n=9

Using the table of random numbers, assign a distinct number to each of the 9 cutting tools in the sample. For example, starting at Row 5, Columns 10-11:

Cutting tool	1	2	3	4	5	6	7	8	9
Random Number	87	20	10	29	66	55	53	49	7

Arrange the random numbers in ascending order:

Cutting tool	9	3	2	4	8	7	6	5	1
Random Number	7	10	20	29	49	53	55	66	87

Treatment 1 (125 in/min): Cutting tools 9, 3, 2
Treatment 2 (150 in/min): Cutting tools 4, 8, 7
Treatment 3 (175 in/min) Cutting tools 6, 5, 1

Example 2

A product development engineer is interested in investigating the tensile strength of a new synthetic fiber that will be used to make cloth for men's shirts. The engineer believes that the strength is affected by the weight percent of cotton used in the blend of materials for the fiber. Furthermore, he suspects that increasing the cotton content will increase the strength. The engineer decides to test specimens at five levels of cotton weight percent: 15%, 20%, 25%, 30%, and 35%. He also decides to test five specimens at each level of cotton content.

The 25 runs should be made in random order. It is very likely that the tensile strength testing machine exhibits a warm-up effect such that the longer it is on, the lower the observed tensile strength readings will be, then the warm-up effect will potentially contaminate the tensile strength data.

Response variable:	Tensile strength
Factor:	Weight percentage of cotton
5 Treatments:	15%, 20%, 25%, 30%, 35%
Experimental unit:	specimens of synthetic fiber

Example 2 (cont'd)

- Using CRD with 5 replicates, n=25

  Using the table of random numbers, assign a distinct number to each of the synthetic fiber in the sample. For example, starting at Row 10, Columns 5-6:

Specimen	1 15%	2 15%	3 15%	4 15%	5 15%	6 20%	7 20%	8 20%	9 20%	10 20%	11 25%	12 25%	13 25%
Random Number	17	88	37	52	09	75	03	12	21	77	66	20	44

Specimen	14 25%	15 25%	16 30%	17 30%	18 30%	19 30%	20 30%	21 35%	22 35%	23 35%	24 35%	25 35%
Random Number	61	32	07	01	82	15	67	04	40	24	99	64

Arrange the random numbers in ascending order:

Specimen	17 30%	7 20%	21 35%	16 30%	5 15%	8 20%	19 30%	1 15%	12 25%	9 20%	23 35%	15 25%	3 15%
Random Number	01	03	04	07	09	12	15	17	20	21	24	32	37

Specimen	22 35%	13 25%	4 15%	14 25%	25 35%	11 25%	20 30%	6 20%	10 20%	18 30%	2 15%	24 35%
Random Number	40	44	52	61	64	66	67	75	77	82	88	99

Advantages and Disadvantages

Advantages:

- Design is flexible. Any number of treatments and replicates may be used in CRD. The number of replications may even vary from treatment to treatment. All the available experimental units may be used in the experiment.
- The statistical analysis is easy even if the number of replications vary or when there are missing data.

Disadvantage:

- Experimental error is expected to be large when the experimental units are not homogeneous.

Randomized Complete Block Design (RCBD)

- *Description:*

  The experimental units are divided into blocks, each of which will constitute a single replication for each treatment. The units belonging in the same block must be homogeneous or closely comparable.

- *Randomization:*

  When the units have been grouped into blocks, the treatments are assigned at random to the units within the same block. A new randomization is made for every group.

Example

Three different machines are being considered in assembling a particular product. The basic measurements on which the machines will be compared is time (in seconds) to complete the assembly. The operation of the machines requires a certain amount of physical dexterity and it is known that there is a difference among the operators in the speed with which they operate the machines. It is decided that 6 different operators are to be used in the experiment.

Response variable:	Speed of assembly (in seconds)
Factor:	Machine model
3 Treatments:	a, b, c
Blocks:	Operators

Example (cont'd)

The order of the three runs of each operator must be randomized. For example, starting at row 0, column 20-21:

Operator	Machine	Random number	Order
1	a	79	3
	b	42	2
	c	13	1
2	a	58	2
	b	92	3
	c	46	1
3	a	16	1
	b	91	3
	c	62	2

Operator	Machine	Random number	Order
4	a	10	1
	b	47	3
	c	45	2
5	a	50	2
	b	89	3
	c	23	1
6	a	59	1
	b	98	3
	c	86	2

Latin Squares (Double Grouping)

- *Description:*

  The design allows for two blocking factors. The treatments are grouped into replicates in two different ways. Every row and every column of any square is a complete replication. The effect of the double grouping is to eliminate from the errors all differences among rows and equally all differences among columns.

- *Randomization:*

  For 5x5 tables, arrange the row and columns of the following table at random:

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

  Note: The Latin letters correspond to the treatments. The number of rows and columns must be the same as the number of treatments. Each treatment occurs only once in each row and once in each column.

Example

The effect of five different ingredients (A,B,C,D,E) on the reaction time of a chemical process is being studied. Each batch of new material is divided into 5 equal parts to permit five runs to be made. Furthermore, each run requires approximately 1.5 hours, so only five runs can be made in one day. The experimenter decides to run the experiment as a latin square so that day and batch effects may be systematically controlled (5 days, 5 batches)

Example (cont'd)

- Arrange rows and columns at random. For example, starting at row 4, columns 0-1,

Row	Random Number	Order	Column	Random Number	Order
1	05	1	1	82	4
2	41	5	2	96	5
3	28	3	3	34	2
4	17	2	4	06	1
5	40	4	5	62	3

- Original square

	1	2	3	4	5
1	A	B	C	D	E
2	B	A	E	C	D
3	C	D	A	E	B
4	D	E	B	A	C
5	E	C	D	B	A

Permuted square

			Day		
Batch	1	2	3	4	5
1	D	C	E	A	B
2	A	B	C	D	E
3	E	A	B	C	D
4	B	D	A	E	C
5	C	E	D	B	A

Exercise

The objective of the experiment is to measure the effectiveness of 5 different fertilizers (A,B,C,D,E) in improving the yield of a certain crop. A piece of agricultural farmland was equally divided into 25 plots as follows:

Plot 1	Plot 2	Plot 3	Plot 4	Plot 5
Plot 6	Plot 7	Plot 8	Plot 9	Plot 10
Plot 11	Plot 12	Plot 13	Plot 14	Plot 15
Plot 16	Plot 17	Plot 18	Plot 19	Plot 20
Plot 21	Plot 22	Plot 23	Plot 24	Plot 25

Exercise (cont'd)

If CRD was used, then we use a randomization mechanism to assign the treatment among all the plots without restriction. For example, assign random numbers starting at row 5, cols 30-31.

Random No.	Plot 1 64	Plot 2 21	Plot 3 39	Plot 4 84	Plot 5 78
Order/Treatment	15 (C)	5 (A)	10 (B)	21 (E)	20 (D)
Random No.	Plot 6 86	Plot 7 60	Plot 8 63	Plot 9 85	Plot 10 02
Order/Treatment	23 (E)	13 (C)	14 (C)	22 (E)	1 (A)
Random No.	Plot 11 08	Plot 12 32	Plot 13 26	Plot 14 50	Plot 15 88
Order/Treatment	2 (A)	8 (B)	7 (B)	11 (C)	24 (E)
Random No.	Plot 16 14	Plot 17 98	Plot 18 74	Plot 19 25	Plot 20 67
Order/Treatment	4 (A)	25 (E)	19 (D)	6 (B)	16 (D)
Random No.	Plot 21 33	Plot 22 11	Plot 23 69	Plot 24 54	Plot 25 73
Order/Treatment	9 (B)	3 (A)	17 (D)	12 (C)	18 (D)

Exercise (cont'd)

If RCBD was used instead so that all the plots in the same row form one block, then we use a randomization mechanism to assign the treatment within a block. For example, assign random numbers starting at row 5, cols 30-31.

Block 1 Random no.	Plot 1 64	Plot 2 21	Plot 3 39	Plot 4 84	Plot 5 78
Order/Treatment	3 (C)	1 (A)	2 (B)	5 (E)	4 (D)
Block 2 Random no.	Plot 6 86	Plot 7 60	Plot 8 63	Plot 9 85	Plot 10 02
Order/Treatment	5 (E)	2 (B)	3 (C)	4 (D)	1 (A)
Block 3 Random no.	Plot 11 08	Plot 12 32	Plot 13 26	Plot 14 86	Plot 15 50
Order/Treatment	1 (A)	3 (C)	2 (B)	5 (E)	4 (D)
Block 4 Random no.	Plot 16 88	Plot 17 14	Plot 18 98	Plot 19 74	Plot 20 25
Order/Treatment	4 (D)	1 (A)	5 (E)	3 (C)	2 (B)
Block 5 Random no.	Plot 21 67	Plot 22 33	Plot 23 11	Plot 24 69	Plot 25 54
Order/Treatment	4 (D)	2 (B)	1 (A)	5 (E)	3 (C)

Exercise (cont'd)

If latin squares was used so that rows of plots serve as the first blocking variable and columns of plots serve as the second blocking variable, then we use a randomization mechanism to permute the rows and columns of the original square. For example, starting at row 5, cols 30-31.

Row	Random Number	Order	Column	Random Number	Order
1	64	3	1	86	5
2	21	1	2	60	2
3	39	2	3	63	3
4	84	5	4	85	4
5	78	4	5	02	1

- Original square

	1	2	3	4	5
1	A	B	C	D	E
2	B	A	E	C	D
3	C	D	A	E	B
4	D	E	B	A	C
5	E	C	D	B	A

Permuted square

			Column		
Row	1	2	3	4	5
1	D	A	E	C	B
2	B	D	A	E	C
3	E	B	C	D	A
4	A	C	D	B	E
5	C	E	B	A	D