HANDOUT # 2 - EXPERIMENTAL DESIGNS

C1. Method of Randomization:

1. Completely Randomized Design (CRD)

2. Randomized Complete Block Design (RCBD)

3. Balanced Incomplete Block Design (BIBD)

4. Latin Square Design (LSD)

5. Crossover Design

6. Split Plot Design

7. Many others

C2. Treatment Structure

1. One Way Classification

2. Factorial - Crossed

3. Factorial - Nested

4. Factorial - Crossed & Nested

5. Fractional Factorial

6. Fixed, Random, Mixed factor levels

C3. Measurement Structure

1. Single measurement on experimental unit

2. Multiple measurements on same EU:

    a. Repeated Measures: One Treatment per EU with Longitudinal or Spatial Indexing of Measurements

    b. Subsampling: One Treatment per EU with Measurements on randomly sampled portions of EU

    c. Crossover Design: Multiple Treatments per EU with Single measurement per treatment

Designed Experiments

In designed experiments, the researcher has control over the important aspects of the study. The control is obtained by either

1. Randomly assigning experimental units (EU's) to treatments or treatment groups:

   a. New drug, standard drug, placebo

   b. Five levels of fertilizer - 100, 200, 300, 400, 500 pound/acre

   c. Three teaching methods - Standard Lecture Course, Web Course with Lectures and Course notes, Web Course with Course Notes but no Lectures

2. Selecting subjects at random from different populations:

   a. Political affiliation - Republican, Democrat, Independent, Other

   b. Type of Elementary School - Public, Private-Nonreligious, Private-Religious

   c. Categories of Lakes - NonPolluted Lakes, Heavily Polluted Lakes, Mildly Polluted Lakes

The two methods lead to similar types of comparisons. The first method is a controlled experiment requiring the researcher to make a random assignment of EU's to the treatments with the EU's selected from a homogeneous population. When this is not possible, important covariates are measured on the EU's.

The second method requires a random sample of EU's from distinct natural populations. In this type of experiment, the researcher controls the random sampling from the different populations but not the assignment of the EU's to the treatment groups.

In nearly all designed experiments, there are extraneous factors which may have an influence on the response. The purpose of the randomization is to yield data such that on the average any major differences in responses between the treatment groups can be attributed to the treatments and not to the extraneous factors for which there is no control.

In all designed experiments, it is crucial that a protocol for the experiment be established prior to the start of the experiment. Included in the protocol are the method by which the EU's are selected, the method of randomization of EU's to the treatment groups, and a plan on how measurements are obtained and recorded. There should be a procedure established to ensure that all decisions in the implementation of the design are made by an objective, repeatable basis.

In observational studies, where there is very little control of extraneous factors, the analysis of the data must be done with great caution. The interpretation of the results from the study often can not be extended beyond the observed data. Inferences about associated populations may be invalid because the researcher does not have control of important features of the study. Therefore, there needs to be a careful discussion of limitations of any statistical inferences from the observed data to the associated populations. For example, in a study of the health differences of individuals having differing amounts fiber in their diet, the factor under study was the level of daily consumption of fiber, Low, Medium, and High. The participants in the study were not randomly assigned to the three treatment groups, but instead were self selected based on a variety of genetic, ethnic, behavioral, and environmental factors which would be unknown to the researcher. Observed differences in the rate of colon cancer in the three groups may be due to factors related to the self-selection and not due to the amount of fiber intake. Methods of analysis for these types of studies are covered in epidemiology, biostatistics, and social science courses.

Components of a Research Design

The major purpose of a designed experiment is to provide the framework for comparing treatment groups in terms of parameters associated with a response variable. The description of a designed experiment consists of three components:

C1: The Method of Randomization

The method by which the EU's are assigned to the Treatments or Subjects are selected from the treatment populations

C2: The Treatment Structure

How the treatments are constructed

C3: The Measurement Process

The method by which the response is recorded on the EU after the EU has been randomly assigned to a treatment or randomly selected from a treatment group

This handout will provide a wide variety of selections and combinations of the three components, including a discussion of some of the advantages and disadvantages of each of the designs.

C2: The Treatment Structure - How the treatments are constructed

The researcher's greatest concern, in most situations, is the treatment structure: The set of factors to be studied and compared using the experimental data.

I. One-way Treatment Structure -

The treatments are t distinct methods of performing a process or t distinct populations:

1. t populations

   $t = 5$ Suppliers of raw materials: S_1, S_2, S_3, S_4, S_5

   $t = 2$ Gender: Female or male

2. t Methods of performing a process

   $t = 3$ Three Methods to teach elementary students how to read

   $t = 4$ Four techniques for inserting chips in a computer board

3. t procedures for treating an illness

   $t = 3$ three Drugs: D_1, D_2, or Placebo

   $t = 2$ two Surgical techniques: New vs Standard

   $t = 5$ five Dose levels of the same drug: 0, 5, 10, 15, 20 mg/100 pounds of subject

II. Factorial Treatment Structure -

The t treatments to be studied in this type of experiment consist of t distinct combinations of the levels of two or more factors. The treatments used in the study may be all possible combinations of the factors or a selected subset. The factors may be crossed or nested.

Definition: Two factors F_1 with a levels and F_2 with b levels are said to be **crossed** if the physical properties of the b levels of F_2 remain the same for all levels of F_1.

Example: An experiment is designed to study the factors which affect the time to fabricate an automobile part from a specimen of metal. Three factors are identified:

1. F_1- Type of Alloy: A, B, C

2. F_2- Porosity of Alloy: .2, .4, .6 mm

3. F_3- Amount of Lubricant Used in Cutting Machine: 0, 5, 10, 15 mL/sec

There would be a total of $t = (3)(3)(4) = 36$ distinct treatments consisting of a level of F_1 combined with a level of F_2 combined with a level of F_3:

Treatments:

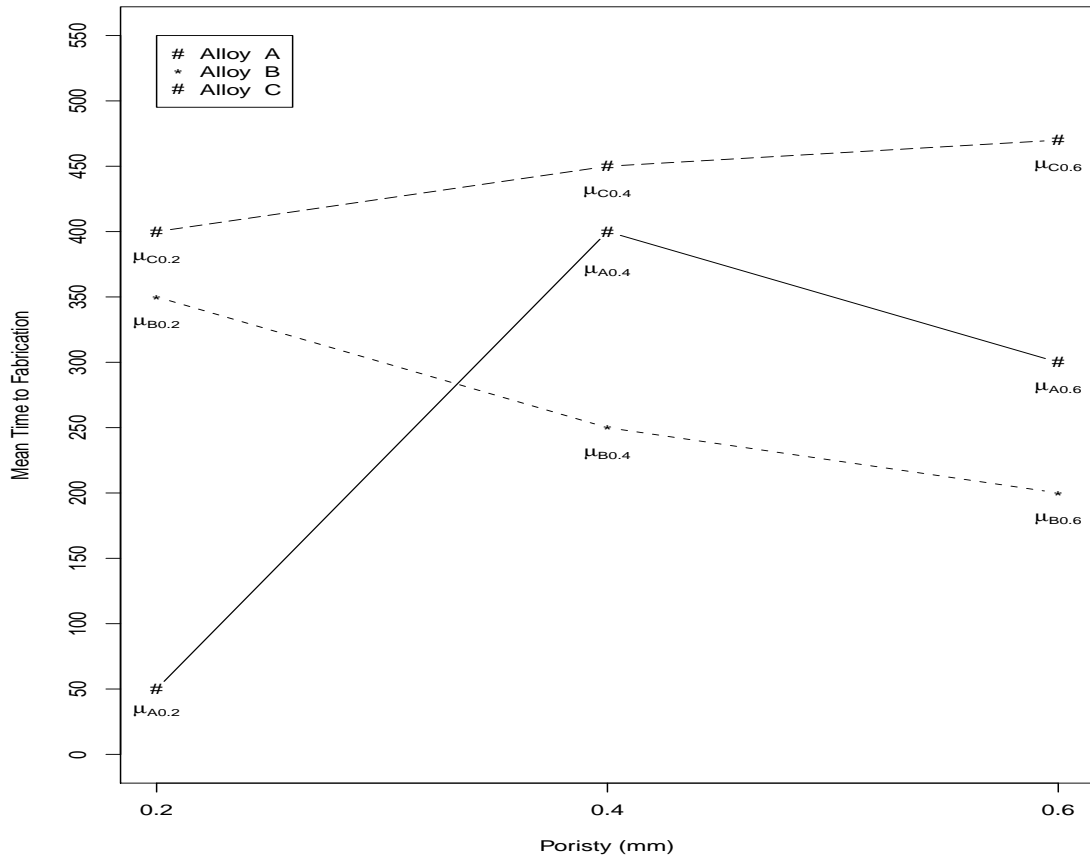|							
(A,.2, 0)	(A,.2, 5)	(A,.2, 10)	(A,.2, 15)	(A,.4, 0)	(A,.4, 5)	(A,.4, 10)	(A,.4, 15)
(A,.6, 0)	(A,.6, 5)	(A,.6, 10)	(A,.6, 15)	(B,.2, 0)	(B,.2, 5)	(B,.2, 10)	(B,.2, 15)
(B,.4, 0)	(B,.4, 5)	(B,.4, 10)	(B,.4, 15)	(B,.6, 0)	(B,.6, 5)	(B,.6, 10)	(B,.6, 15)
(C,.2, 0)	(C,.2, 5)	(C,.2, 10)	(C,.2, 15)	(C,.4, 0)	(C,.4, 5)	(C,.4, 10)	(C,.4, 15)
(C,.6, 0)	(C,.6, 5)	(C,.6, 10)	(C,.6, 15)				

A complete $3 \times 3 \times 4$ factorial experiment with all 3 factors **crossed** would have the levels of F_2 exactly the same for all levels of both factors F_1 and F_3. Also, the levels of F_3 would be the same for all levels of F_1 and F_2. That is, the three Porosity levels would be the same for all three types of Alloys and the amount of Lubricant would be the same for all combinations of Type of Alloy and Porosity.

One of the major goals in an experiment consisting of two or more factors is to determine if the effects of the factors interact. That is, is the size of the differences in the mean responses of two levels of one of the factors the same or different across the various levels of a second factor? For example, suppose we have two factors F_1 and F_2 with 3 and 4 levels, respectively. If the two factors do not interact then the mean responses must satisfy:

$$\mu_{ij} - \mu_{kj} = \mu_{ih} - \mu_{kh} \ \text{ for } \ i \neq k = 1, 2, 3 \ \text{ and } \ j \neq h = 1, 2, 3, 4$$

If the factors interact, then there exist a set of treatments for which the above equality does not hold. The goal of the study is test the above hypotheses and then measure the size of the effects $\mu_{ij} - \mu_{kj}$ if an interaction exists. Graphically, we would have the following profile plots for the various situations in which interaction exists or not:

Profile Plot of Alloy*Porosity Interaction

Mean Time to Fabrication

Alloy A
* Alloy B
# Alloy C

$\mu_{C0.2}$ $\mu_{C0.4}$ $\mu_{C0.6}$

$\mu_{B0.2}$ $\mu_{A0.4}$ $\mu_{A0.6}$

$\mu_{B0.4}$ $\mu_{B0.6}$

$\mu_{A0.2}$

Poristy (mm)

Profile Plot of Alloy*Porosity Interaction

Mean Time to Fabrication

Alloy A
* Alloy B
# Alloy C

$\mu_{A0.6}$

$\mu_{A0.4}$

$\mu_{B0.6}$

$\mu_{A0.2}$ $\mu_{B0.4}$ $\mu_{C0.6}$

$\mu_{C0.4}$

$\mu_{B0.2}$

$\mu_{C0.2}$

6

Poristy (mm)
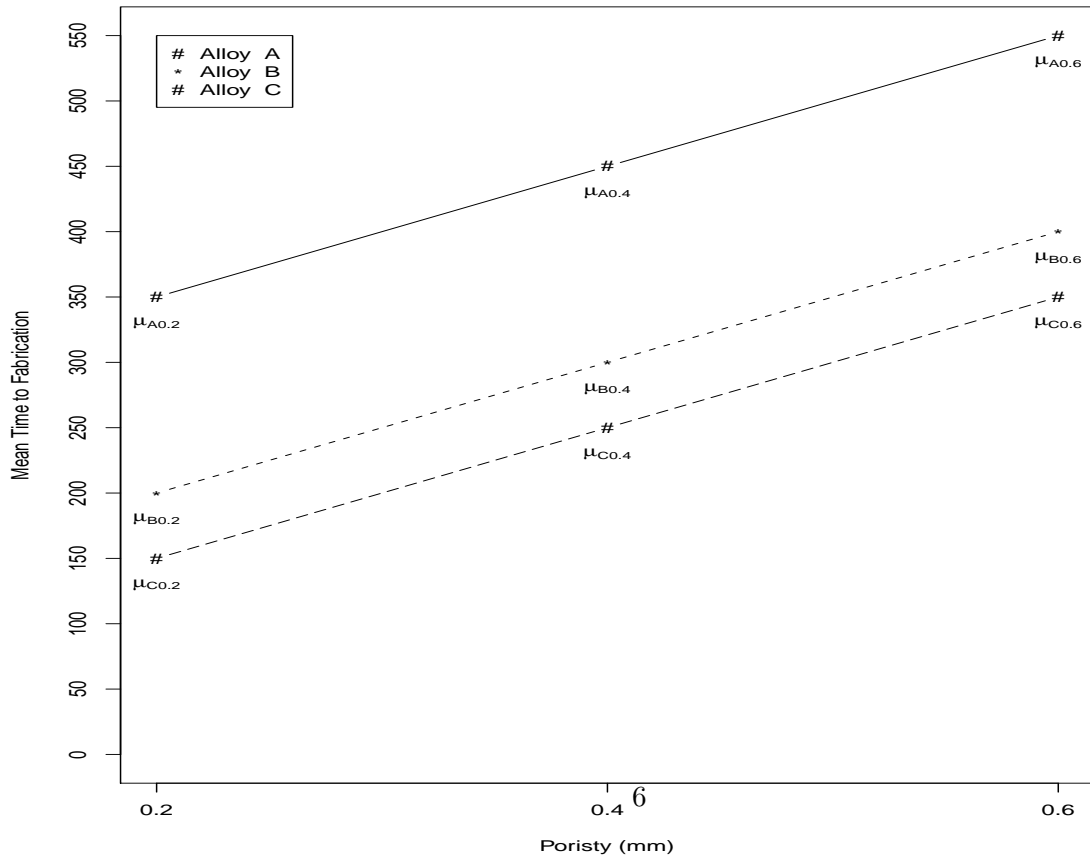
Profile Plot of Alloy*Porosity Interaction
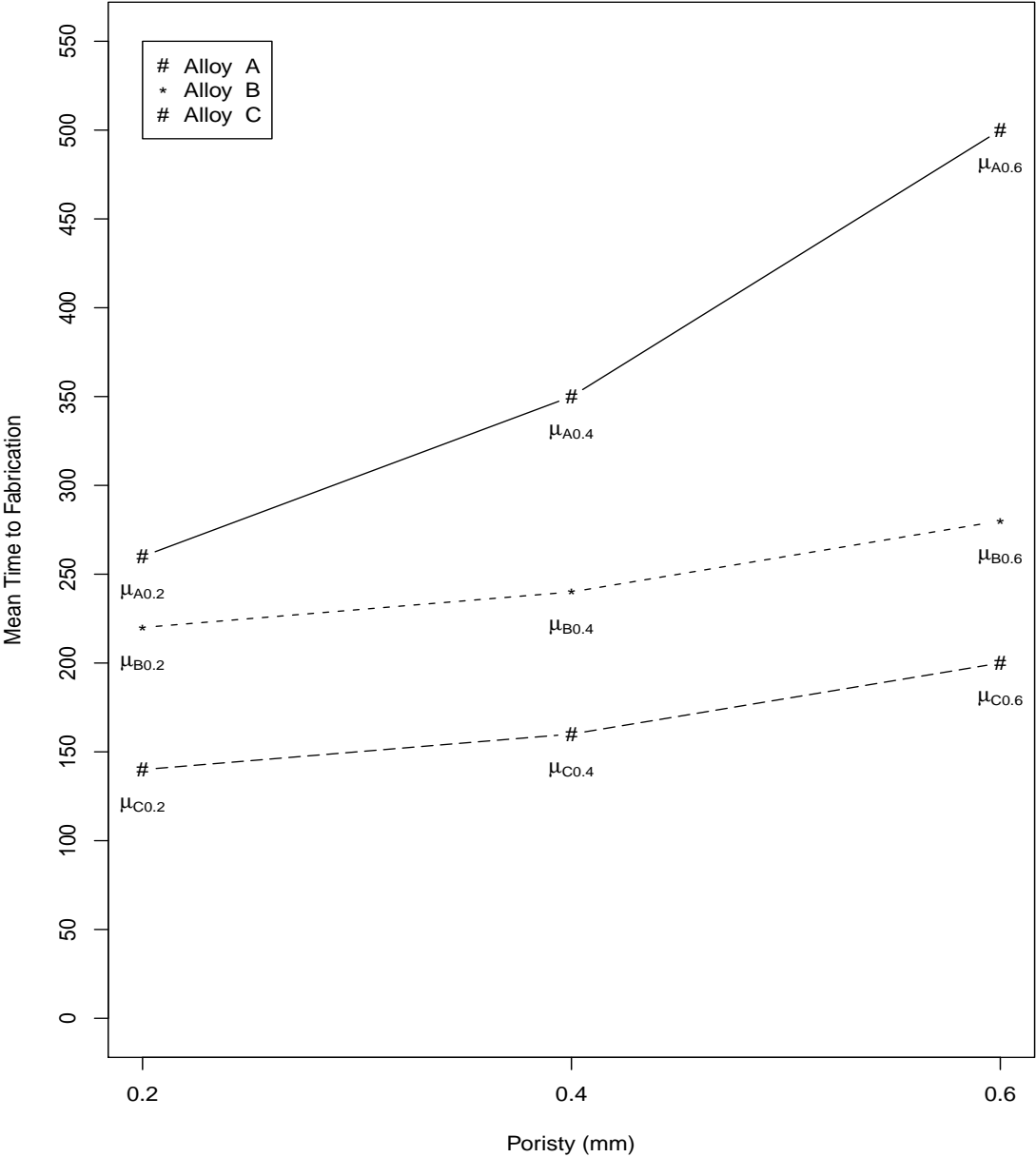
Nested Treatment Structure -

In some multifactor experiments, the levels of one factor (e.g., factor B) are not identical for different levels of a second factor (e.g., factor A). Such an arrangement of factors is called a **nested**, or **hierarchical**, treatment structure, with the levels of factor B nested within the levels of factor A.

Definition: The levels of Factor F_2 are said to be **nested within** the levels of factor F_1 if the physical properties of the levels of factor F_2 vary depending on which level of factor F_1 is used.

Example: For many Alloys, the possible porosity levels may vary. For example, the porosity level of three alloys are classified as Low, Medium, and High. However, the definition of Low, Medium, and High is quite different for the three alloys. The true porosity levels in g/cm^3 for the three alloys are given in the following table.

| | Porosity Level | | |
Alloy Type	Low	Medium	High
A1	.2	.3	.4
A2	.3	.5	.7
A3	.5	.8	.95

Thus, the definition of level (Low, Medium, High) of the factor Porosity differ depending upon which level of the factor Type of Alloy is being considered. That is, the levels of factor F_2, Porosity are Nested Within the levels of factor F_1, Type of Alloy.

We denote factor F_2 being nested within factor F_1 as $F_2(F_1)$.

Example: A company is designing a new product. There are 3 producers of raw material. The goal of a study was to evaluate the consistency of the raw material's physical properties across multiple batches of material from each producer and then study the variation within each batch. The experiment consisted of randomly select 6 batches of material from each of the 3 producers. From each of the 18 batches, 4 samples are taken and the physical properties were then determined in the company's lab. The factors are

Producer (P): P1, P2, P3

Batch (B): B1, B2, B3, B4, B5, B6

Sample (S): S1, S2, S3, S4

The levels of Batch only having meaning if we know from which Producer the batches are selected. Thus, the factor Batch is nested within the levels of factor Producer: B(P).

The levels of Sample only having meaning if we know from which producer and batch the sample is selected. Thus, the factor Sample is nested within both the levels of factor Producer and the factor Batch: S(B,P).

This would be designated as a hierarchical nested experiment with factors:

P, B(P), S(B,P)

Example: A social scientist is studying the use of drugs by high school students. The factors of interest are Type of High School (private-nonreligious, private-religious,public); Particular Schools within each Type of School; Individual Students from each School within each Type of School.

Are the factors crossed or nested.

Example: The filling machine for a soft drink bottler theoretically fills each bottle to the correct target height. However, in practice, there is deviation about the target value and the process engineer designs a study to investigate how to achieve greater uniformity. There are three variables which can be controlled during the filling process:

F_1 :   Percent Carbonation - 10%, 12%, 14%

F_2 :   Operating Pressure - 25 psi, 30 psi

F_3 :   Line Speed - 200 bpm, 250 bpm

Are the factors crossed or nested.

Example: An animal scientist is studying the level of infestation of ticks on cattle. The factors are

F_1 :   Breed of Cattle - Angus, Beefalo, Limousin, Salers

F_2 :   Individual cows of each Breed - Cow1, Cow2, ..., Cow25

F_3 :   Locations on cow - Head, Neck, Flank, Hip, Tail

Are the factors crossed or nested

III. Fractional factorial

Fractional factorial treatment structures are important alternatives to complete factorial treatment structures when cost, time, or experimental constraints preclude the execution of complete factorial design. In addition, experiments that involve many factors are routinely conducted as fractional factorials because it is not necessary to test all possible factor-level combinations in order to estimate the most crucial factor effects, main effects and low-order interactions.

An experiment was designed to study the corrosion rate of a reactor at a chemical acid plant. There were five factors of interest: each having two levels.

Factor Levels for Acid-Plant Corrosion-Rate Study

Factor	Levels
A_1: Raw-material feed rate	3000, 6000 pph
A_2: Gas temperature	$100°C$, $220 °C$
A_3: Scrubber water	5%, 20%
A_4: Reactor-bed acid	20%, 30%
A_5: Exit temperature	$300°C$, $360°C$

Initially it was unknown whether changing any of these five process variables (factors) would have an effect on the reactor corrosion rate, the response variable of interest to the process engineers. In order to minimize maintenance and downtime, it was desirable to operate this acid plant under process conditions which produce a low corrosion rate. If one were to test all possible combinations of three or four levels of the five process variables, 3^5=243 or 4^5=1024 runs would be needed for a single replication of the possible treatments. A practical difficulty with this investigation was that the plant needed to cease commercial production for the duration of the study. An experimental design that required 243 or 1024 runs would have been prohibitively expensive. A small screening experiment was designed in which only gross factor effects were identified, the feasible extremes of the five variables. It is still necessary to execute $2^5 = 32$ runs. However, the cost of experimentation would be greatly reduced over an experiment involving three or four levels of each of the factors.

The table on the next page displays this proposed design using the notation: -1 for the low level of the factor and $+1$ for the high level of the factor.

An alternative notation designates the 32 treatments: If a factor is at its low level then its symbol is not displayed. Thus, we have that

Treatment $a_1a_2a_5$ represents Factors

A_1, A_2, A_5 at their High Level; and Factors A_3, A_4 at their Low Level

The treatment having all Factors at their low level is designated as I or as (1).

The following table illustrates this notation for all $2^5 = 32$ treatments:

Table 1: Treatments 2^5 Factorial

TRT	A_1	A_2	A_3	A_4	A_5	TREATMENT	Y_{ijklm}
1	-1	-1	-1	-1	-1	(I)	Y_{11111}
2	-1	-1	-1	-1	1	a_5	Y_{11112}
3	-1	-1	-1	1	-1	a_4	Y_{11121}
4	-1	-1	-1	1	1	a_4a_5	Y_{11122}
5	-1	-1	1	-1	-1	a_3	Y_{11211}
6	-1	-1	1	-1	1	a_3a_5	Y_{11212}
7	-1	-1	1	1	-1	a_3a_4	Y_{11221}
8	-1	-1	1	1	1	$a_3a_4a_5$	Y_{11222}
9	-1	1	-1	-1	-1	a_2	Y_{12111}
10	-1	1	-1	-1	1	a_2a_5	Y_{12112}
11	-1	1	-1	1	-1	a_2a_4	Y_{12121}
12	-1	1	-1	1	1	$a_2a_4a_5$	Y_{12122}
13	-1	1	1	-1	-1	a_2a_3	Y_{12211}
14	-1	1	1	-1	1	$a_2a_3a_5$	Y_{12212}
15	-1	1	1	1	-1	$a_2a_3a_4$	Y_{12221}
16	-1	1	1	1	1	$a_2a_3a_4a_5$	Y_{12222}
17	1	-1	-1	-1	-1	a_1	Y_{21111}
18	1	-1	-1	-1	1	a_1a_5	Y_{21112}
19	1	-1	-1	1	-1	a_1a_4	Y_{21121}
20	1	-1	-1	1	1	$a_1a_4a_5$	Y_{21122}
21	1	-1	1	-1	-1	a_1a_3	Y_{21211}
22	1	-1	1	-1	1	$a_1a_3a_5$	Y_{21212}
23	1	-1	1	1	-1	$a_1a_3a_4$	Y_{21221}
24	1	-1	1	1	1	$a_1a_3a_4a_5$	Y_{21222}
25	1	1	-1	-1	-1	a_1a_2	Y_{22111}
26	1	1	-1	-1	1	$a_1a_2a_5$	Y_{22112}
27	1	1	-1	1	-1	$a_1a_2a_4$	Y_{22121}
28	1	1	-1	1	1	$a_1a_2a_4a_5$	Y_{22122}
29	1	1	1	-1	-1	$a_1a_2a_3$	Y_{22211}
30	1	1	1	-1	1	$a_1a_2a_3a_5$	Y_{22212}
31	1	1	1	1	-1	$a_1a_2a_3a_4$	Y_{22221}
32	1	1	1	1	1	$a_1a_2a_3a_4a_5$	Y_{22222}

The goal in the design of fractional factorial experiments is to ensure that the effects of primary interest are either unconfounded with other effects or, if that is not possible, confounded with effects that are not likely to have appreciable magnitudes. Confounding was defined as the situation where an effect cannot unambiguously be attributed to a single main effect or interaction.

In general, when designing fractional factorial experiments one seeks to confound either effects known to be negligible relative to the uncontrolled experimental error variation, or in the absence of such knowledge, high-order interactions, usually those involving three or more factors. Confounding of high-order interactions is recommended because frequently these interactions either do not exist or are negligible relative to main effects and low-order interactions. Thus, in the absence of the information provided by the pilot-plant study, it would be preferable to confound the three-factor interaction in the manufacturing-plant experiment rather than the main effect for temperature.

Not only does confounding occur when a complete factorial experiment is conducted in blocks; it also occurs when only a portion of all the possible factor-level combinations are included in the design. Confounding occurs because two or more effects representations (apart from a change in all the signs) are the same. A calculated effect then represents the combined influence of the effects. In some instances (e.g., one-factor-at-a-time experiments) the confounding pattern may be so complex that one cannot state with assurance that any of the calculated effects measure the desired factor effects. This again is why planned confounding, confounding in which important effects either are unconfounded or are only confounded with effects that are believed to be negligible, is the basis for the statistical constructions of fractional factorial experiments.

In our example, suppose the company only can afford to have 16 of the necessary 32 runs conducted. The following design has no main effect confounded with any other main effect or with any two-way interaction but some of the two-way interactions are confounded with each other. Thus, the main effects can all be estimated provided we assume that three-way interactions or higher are negligible. In Handout 10, we will discuss how to select which treatments to include in the experiment so the main effects and two-way interactions are appropriately confounded with higher order interactions.

Table 2: A 1/2 fraction of a 2^5 factorial experiment (2^{5-1})

TRT	A_1	A_2	A_3	A_4	A_5	$A_1A_2A_3A_4A_5$
2	-1	-1	-1	-1	1	1
3	-1	-1	-1	1	-1	1
5	-1	-1	1	-1	-1	1
8	-1	-1	1	1	1	1
9	-1	1	-1	-1	-1	1
12	-1	1	-1	1	1	1
14	-1	1	1	-1	1	1
15	-1	1	1	1	-1	1
17	1	-1	-1	-1	-1	1
20	1	-1	-1	1	1	1
22	1	-1	1	-1	1	1
23	1	-1	1	1	-1	1
26	1	1	-1	-1	1	1
27	1	1	-1	1	-1	1
29	1	1	1	-1	-1	1
32	1	1	1	1	1	1

The following table contain a 1/4 fraction of the 32 treatments which would thus require only 8 runs to complete the experiment. Of course this design is not very realistic in that even the main effects are confounded.

Table 3: A 1/4 fraction of a 2^5 factorial experiment: 2^{5-2}

Using $A_1A_2A_3$ and $A_1A_2A_3A_4A_5$ as the Generators

TRT	A_1	A_2	A_3	A_4	A_5	$A_1A_2A_3$	$A_1A_2A_3A_4A_5$	A_4A_5
1	-1	-1	-1	-1	-1	-1	-1	1
4	-1	-1	-1	1	1	-1	-1	1
13	-1	1	1	-1	-1	-1	-1	1
16	-1	1	1	1	1	-1	-1	1
21	1	-1	1	-1	-1	-1	-1	1
24	1	-1	1	1	1	-1	-1	1
25	1	1	-1	-1	-1	-1	-1	1
28	1	1	-1	1	1	-1	-1	1

Notice that the coefficients for A_4 and A_5 are identical and hence the main effects of factors A_4 and A_5 are confounded.

IV. Random Levels of Factor(s)

Thus far we have discussed experiments in which the levels of the Factors are preselected by the researcher and these are the only levels of the factors of interest to the researcher. In some experiments, the levels of the factor(s) may be randomly selected from a population of possible levels. The researcher is interested in making inferences about all levels but only has the time and/or resources to examine a subset of these levels.

Definition   A factor has **random effects** if the levels of the factor to be used in the experiment or study are randomly selected from a population of potential levels.

The researcher is interested in all levels of the factor but because of time or economical constraints it is not possible to examine all possible levels. Thus, the responses are observed on a possibly small subset of all possible factor levels (treatments) but inferences are to be about all the levels (treatments).

EXAMPLE 1:   A survey is to be designed to estimate the variability in yield per acre of cotton across farms in Texas. Twenty counties are randomly selected from the 254 counties in Texas. Next, several farms are randomly selected within each county. At each farm, several plots of 1-acre each are randomly selected to determine the yield. The individual measurements of yield on each plot would depend on the variability across counties and the variability across farms within a county. This type of experimental design provides inferences which are considerably different from the inferences that would be made to compare cotton production at a few designated counties within Texas.

- C: Counties - C_1, C_2, ..., C_{20} have random levels

- F: Farms -

    F_1, F_2, ..., F_{n_1} - n_1 Farms from County C_1

    F_{n_1+1}, F_{n_1+2}, ..., $F_{n_1+n_2}$ - n_2 Farms from County C_2

    $\vdots$

    $F_{n_1+\cdots+n_{19}+1}$, $F_{n_1+\cdots+n_{19}+2}$, ..., $F_{n_1+\cdots+n_{19}+n_{20}}$ - n_{20} Farms from County C_{20}

    The levels of Farms are random and nested within the levels of Counties,

    denoted by $F_j(C_i)$.

- P: Plots - At each Farm within each County $F_j(C_i)$, there are $m_{i,j}$ Plots which are randomly selected within $F_j(C_i)$

    The levels of Plots are random and nested within the levels of Farms within Counties,

    denoted by $P_k(F_j,\ C_i)$.

-

    In this experiment, the Experimental Unit would be the Plots and the two factors of interest would be Counties and Farms within Counties.

EXAMPLE 2: A company produces fiber optic cable for the communication industry. This industry needs cable which has a very consistent tension across long lengths of cable. The company is planning on purchasing new machinery used to bind the cable. There are three major manufacturers of the machines. The fiber optic company wants to determine if there is major differences in the quality of the product from the three manufacturers. A study is designed by randomly selecting 5 binding machines from each each of the manufacturers and then producing 10 cables of length 1000 yards from each of the 15 machines. The quality of the signal transmitted through the cable is then measured. Research Question: Is there a difference in quality due to

1. Manufacturer Differences

2. Variability in Machines From the Same Manufacturer

3. Variability in Cable Produced From the Same Machine

- M: Manufactures - M_1, M_2, M_3 has fixed levels

- B: Binding Machines -

    B_1, B_2, ..., B_5 - 5 Binding Machines randomly selected from Manufacturer M_1

    B_6, B_7, ..., B_{10} - 5 Binding Machines randomly selected from Manufacturer M_2

    B_{11}, B_{12}, ..., B_{15} - 5 Binding Machines randomly selected from Manufacturer M_3

    The levels of Binding Machines are random and nested within the levels of Manufacturer,

    denoted $B_j(M_i)$.

- C: Cable - For each Binding Machine from each Manufacturer $B_j(M_i)$, 10 cables are randomly selected from the output

    The levels of Cable are random and nested within the levels of Binding Machines within Manufacturer,

    denoted $C_k(B_j, M_i)$.

- In this experiment, the Experimental Unit would be the Cable and the two factors of interest would be Manufacturer and Binding Machine within Manufacturer.

EXAMPLE 3: There are hundreds of laboratories that are federally qualified to produce assessments of the level of *e. coli* in meat. A consumer research group wants to determine if there is a difference in the accuracy of the determinations across the many laboratories or are the laboratories all essentially the same in the quality of the determinations. A study is designed by

1. Randomly selecting 10 laboratories from the list of all laboratories.

2. Randomly assigning 20 meat samples with a known level of *e. coli* to each of the selected laboratories.

3. Measure the difference between the known value of *e. coli* and the determination made by the laboratory.

- What are the factors? Which factors are fixed? random? What is the nesting structure?

How would the inferences differ in the three studies from the possible inferences conducted if all the factors have fixed levels?

In Example 1, the researchers are interested in all counties in Texas but will base their inferences on a sample of counties, farms, and plots within these farms.

In Example 2, the company would like to have information about all machines at each of the three Manufacturers but will make inferences based on a limited sample of such machines. If they find large differences in the selected machines, then they may decide to expand the sample to include a larger proportion of the population of machines at each of the three manufactures.

In Example 3, the researchers are interested in the variability across the population of Laboratories. If there is very little difference across the selected Laboratories, then the consumer group may be willing to conclude that government certification has resulted in a homogeneous set of laboratories with very little difference in their measurement error. Whereas, if large differences are found, they may expand their study to determine why the laboratories are producing such a range of differences in the accuracy of their determination of the level of *e. coli* and to determine if these types of differences persist across the population of all such laboratories.

C1: Method of Randomization

Completely Randomized Design (CRD)

In a CRD, the n experimental units are considered to have no identifiable differences relative to the response. Thus, there is no restrictions on the randomization. For $i = 1, 2, \cdots, t$, the n_i EU's assigned to Treatment # i are n_i randomly selected EU's from the totality of n EU's.

Randomization Procedure

In a CRD with t treatments and n_i EU's per treatment, the procedure for assignment of EU's to treatments can be conducted as follows:

Type I: For Randomized Treatments in a Designed Experiment

Suppose we have t treatments: $A_1, \cdots, A_n$ and n EU's

The $n = n_1 + n_2 + \cdots + n_t$ EU's are considered homogenous relative to characteristics which may affect the response after the treatments have been applied.

1. Assign the numbers 1 to n to the experimental units

2. Obtain a random permutation of the numbers 1 to n

   R code: $y < - $ **sample(n,replace=FALSE)**

3. Assign the first n_1 EU's in the list to Treatment A_1,   the next n_2 EU's in the list to Treatment A_2,   ..., the last n_t EU's to treatment A_t.

Example: Suppose $n = 20$, $t = 3$, $n_1 = 7$, $n_2 = 5$, $n_3 = 8$:

1. $EU_1, EU_2, EU_3, EU_4, EU_5, EU_6, EU_7, EU_8, EU_9, EU_{10},$

   $EU_{11}, EU_{12}, EU_{13}, EU_{14}, EU_{15}, EU_{16}, EU_{17}, EU_{18}, EU_{19}, EU_{20}$

2. A random permutation is obtained using the R code: **sample(20,replace=FALSE)**

   1   15   18   5   4   17   19   3   13   10   9   14   8   12   20   11   16   6   7   2

3. Create 3 groups of 7, 5, and 8 numbers, respectfully :

   <u>1 15 18 5 4 17 19</u>   <u>3 13 10 9 14</u>   <u>8 12 20 11 16 6 7 2</u>

4. The resulting assignment of EU's to treatments is given by

   $A_1 :$ $EU_1,$ $EU_{15},$ $EU_{18},$ $EU_5,$ $EU_4,$ $EU_{17},$ EU_{19}

   $A_2 :$ $EU_3,$ $EU_{13},$ $EU_{10},$ $EU_9,$ EU_{14}

   $A_3 :$ $EU_8,$ $EU_{12},$ $EU_{20},$ $EU_{11},$ $EU_{16},$ $EU_6,$ $EU_7,$ EU_2

Type II: Comparative Study

Suppose we have t existing populations that we want to compare. For example,

1. five manufacturers of fire alarm devices

2. six producers of a raw material in the manufacturer of paint

3. five types of landscapes in which birds may migrate.

The procedure is to take a simple randomly sample of n_i items (EU's) from Population i. That is, select the units such that every unit in the population has an equal chance of being selected. Then make observations or measurements on the $n = n_1 + \cdots + n_t$ selected items.

The difference from a comparative study and a purely observational study is that in the comparative study we randomly select units from existing populations and then make measurements/observations on the selected units. We then use this information (data) to draw inferences about the population.

In a purely observational study, we do not randomly select units from the population to make measurements but take existing data sets which have been complied in some non-random form.

1. A list of parents at an elementary school

2. The medical records of dogs at the county lost animal facility

3. A description of returned items at a department store

4. The students enrolled in STAT 303

The above are what we describe as "convenience" samples from the population. This type of data does not allow us to make valid inferences from the populations from the data was obtained. Why?

Blocking Designs

In the completely randomized design, with either a single factor, factorial, or nested treatment structure, the measurement design consisted of either a single response on each EU or subsampling. In order to conduct a CRD, the EU's were required to be homogeneous relative to their potential effect on the response variable. Also, the conditions under which the experiment was conducted for each replication needed to be essentially identical relative to their potential impact on the response from the EU's. For example, if we were comparing the effectiveness of three drugs on the treatment of a disease, the patients would need to be of the same relative health and with the same severity of the disease. If a greenhouse experiment was conducted to assess the difference in the impact of five growth stimulants on plant growth, it would be necessary to have plants of the same age, size, variety, etc. and for the conditions in the greenhouse to remain relatively the same over the whole experiment.

When the available EU's have differences prior to conducting the experiment or the experimental conditions may vary over the days in which the experiment is conducted, it is important to incorporate this information in the design of the experiment. The characteristics of the EU's which make them different and which may impact on the responses from the EU's can be used to identify groups of EU's such that the EU's within a group will have homogeneous impact on the response variable, whereas EU's in different groups may provide extremely different impact on the responses.

Definition: A **block** is a group of EU's for which the EU's within the block have relatively similar characteristics relative to those characteristics which may impact the response variable. The EU's within a block are much more alike in comparison to EU's in other blocks.

A **complete block** is a homogeneous group of EU's in which the number of EU's per treatment level is the same for all t treatments. In many cases, there are t EU's per block, therefore, each treatment appears only once in each block.

The design of blocked experiments vary depending on whether the blocks of EU's are complete or not. In many situations, there may be experimental constraints which prevent the blocks from being complete. Consider the following example.

Four Designs for Evaluating Tread Wear of Tires

A company wants to evaluate a number of different brands of tires, B_1, B_2, B_3, B_4 with respect to their wear characteristics. The tires will be placed on cars and driven for 30,000 miles on a 2-mile oval track. The loss in tread thickness will be recorded for each tire. Four designs are under consideration for this study.

DESIGN I: Completely Randomized Design (CRD) with 4 Treatments (Brand) and 4 Reps/Treatment:

There are four cars available for the study of four brands. Four tires of each brand are obtained from the manufacturer and the 16 tires are randomly assigned to the four cars. One such random assignment is given below. What are some of the possible problems with using this design?

	CAR			
POSITION	C1	C2	C3	C4
RIGHT FRONT	B4	B1	B3	B1
RIGHT REAR	B1	B2	B1	B3
LEFT FRONT	B4	B3	B3	B2
LEFT REAR	B4	B4	B2	B2

Source of Variation	Degrees of Freedom
Brand	3 (4-1)
Error	12 (15-3)
Total	15 (16-1)

DESIGN II: Randomized Complete Block Design (RCBD) with 4 Treatments (Brand) and one blocking variable (CARS):

Same situation as in Design I except in this design the tires are blocked by car so that one tire of each brand is on each car. Then the tires are randomly assigned to the position on a car. What are some of the improvements and possible problems with using this design?

	CAR			
POSITION	C1	C2	C3	C4
RIGHT FRONT	B1	B1	B1	B1
RIGHT REAR	B3	B4	B2	B4
LEFT FRONT	B4	B2	B3	B2
LEFT REAR	B2	B3	B4	B3

Source of Variation	Degrees of Freedom
Cars	3 (4-1)
Brand	3 (4-1)
Error	9 (15-6)
Total	15 (16-1)

DESIGN III: Latin Square Design (LSD) with 4 Treatments (Brand) and 2 Blocking Variables (CARS) and (POSITIONS on Cars):

In this design the tires are blocked based both on the type of car and the position on car. The randomization is done with respect to the individual tires. What are some of the improvements and possible problems with using this design?

	CAR			
POSITION	C1	C2	C3	C4
RIGHT FRONT	B1	B2	B3	B4
RIGHT REAR	B2	B3	B4	B1
LEFT FRONT	B3	B4	B1	B2
LEFT REAR	B4	B1	B2	B3

Source of Variation	Degrees of Freedom
Cars	3 (4-1)
Position	3 (4-1)
Brand	3 (4-1)
Error	6 (15-9)
Total	15 (16-1)

DESIGN IV: Balanced Incomplete Block Design (BIBD) with 7 Treatments (Brand), 1 Blocking Variable (CARS), and 4 Reps:

In this study there are 7 brands of tires so that not all brands may appear on all cars. However, a somewhat "balanced design" is achieved. What are some of the possible problems with using this design?

	CAR						
POSITION	C1	C2	C3	C4	C5	C6	C7
RIGHT FRONT	B4	B2	B6	B3	B7	B5	B1
RIGHT REAR	B2	B3	B7	B4	B1	B6	B5
LEFT FRONT	B3	B5	B1	B7	B4	B2	B6
LEFT REAR	B1	B7	B2	B6	B5	B4	B3

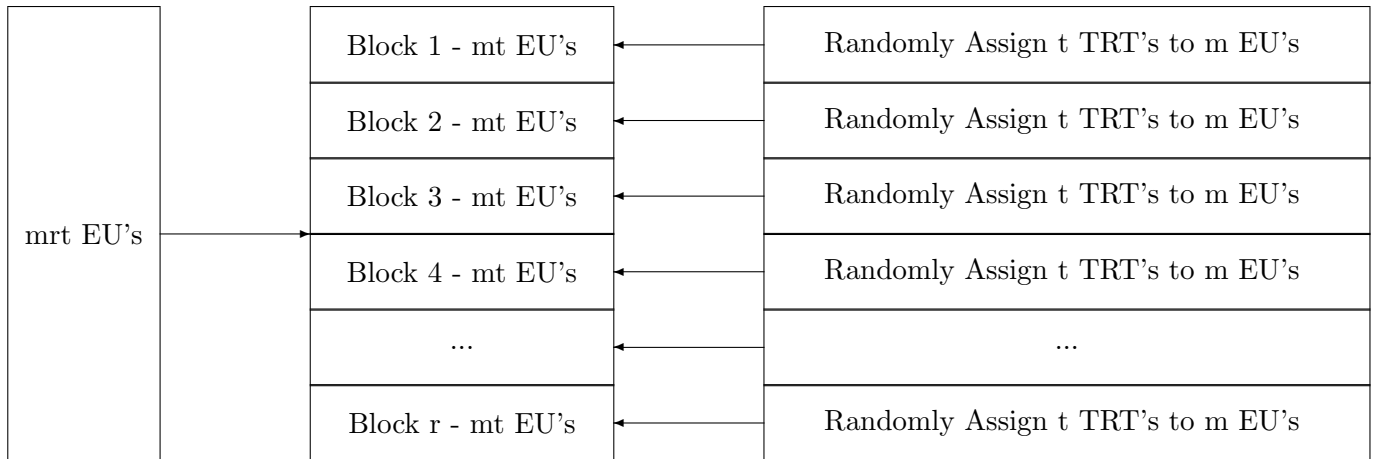| Source of Variation	Degrees of Freedom
Car	6 (7-1)
Position	3 (4-1)
Error	18 (27-9)
Total	27 (28-1)

We will next discuss the advantages and disadvantages of each of the four designs.

RANDOMIZED COMPLETE BLOCK DESIGN (RCBD)

The general setting for a RCBD consists of an experiment with t treatments which may be constructed as a factorial or simply a single factor. There are n EU's which are non-homogeneous or the experimental conditions can not be controlled to be exactly the same for each replication of the experiment. For example, it may take several days to conduct the experiment and there is day to day variation in the conditions in a lab or greenhouse. In the case of non-homogeneous EU's the randomization will involve two stages:

Stage I: The EU's are grouped (or randomly assigned) into r blocks of mt EU's, in many cases $m = 1$. The grouping is done in such a manner that the EU's in the same block have similar characteristics relative to those characteristics which may affect the response from the EU, whereas EU's in different blocks may have very different characteristics.

Stage II: The mt EU's in a given block are then randomly assigned to the t treatments such that m EU's are assigned to each of the t treatments. That is, there is essentially a CRD with m reps per treatment within each of the r blocks.

mrt EU's	Block 1 - mt EU's	Randomly Assign t TRT's to m EU's
	Block 2 - mt EU's	Randomly Assign t TRT's to m EU's
	Block 3 - mt EU's	Randomly Assign t TRT's to m EU's
	Block 4 - mt EU's	Randomly Assign t TRT's to m EU's
	...	...
	Block r - mt EU's	Randomly Assign t TRT's to m EU's

The RCBD sounds like such a good method for designing an experiment and it is nearly impossible to find EU's which are homogeneous, why not use a RCBD in all experiments? The reason is a loss in precision when the blocking is not really needed. That is, if we use r blocks, there is a loss of $r-1$ df from the df for MSE. Thus, if the arrangement of the EU's into blocks does not provide a correspondingly large reduction in the SSE, then the F-test for treatment effects and multiple comparison procedures will have a reduction in precision relative to the corresponding CRD. A measure of the effectiveness of blocking is given by the size of the mean square for the Block Effect relative to MSE in the blocked experiment.

Advantages of a RCBD to a CRD

1. The analysis is relatively direct. A meaningful analysis can be conducted even when some of the observations are missing.

2. A more precise analysis can be obtained from the RCBD compared to the CRD. When the blocking effect is significant, the reduction in MSE affords

   - a more powerful F-test for treatment differences
   - a more precise estimation of the differences in treatment means

3. The RCBD is a very flexible design procedure. There is no limitation, within the design structure, with respect to the number of treatments or the number of blocks. However, there may be a limitation on the number of treatments in the experiment, t, if a block can only contain m EU's with $m < t$.

Disadvantages of a RCBD to a CRD

1. If the number of treatments t is large, it may be difficult to obtain a homogeneous grouping of the EU's. The greater the number of EU's required per block, the greater the chance that the EU's will not be homogeneous within each of the blocks.

2. If there is a block by treatment interaction and $m = 1$, then there is not in general a valid analysis of treatment effects. Under special circumstances, a partial analysis is feasible.

EXAMPLE   The experiment was conducted to compare four different pre-planting treatments for soybean seeds, T_1, T_2, T_3, T_4. A fifth treatment, C, consisting of not treating the seeds, was used as a control. the experimental area consisted of four fields, F_1, F_2, F_3, F_4. Each field was divided into five plots, P_1, P_2, P_3, P_4, P_5 with one of the five treatments randomly assigned to each of the 5 plots within each field. A new randomization was conducted for each of the four fields. The measurements y_{ij} are the number of plants which failed to emerge out of 450 seeds planted per plot.

	Plot				
	P_1	P_2	P_3	P_4	P_5
Field 1	C	T_3	T_1	T_4	T_2
Field 2	T_1	C	T_4	T_2	T_3
Field 3	T_4	T_1	C	T_2	T_3
Field 4	T_2	T_3	C	T_4	T_1

Source of Variation	Degrees of Freedom
Field	3 (4-1)
Treatment	4 (5-1)
Error	12 (19-7)
Total	19 (20-1)

LATIN SQUARE DESIGN (LSD) EXAMPLE

An experiment was conducted to assess the relative resistance to abrasion of four grades of leather (G1, G2, G3, G4). A machine was used in which the samples of leather could be tested in any one of four machine positions. Since different runs (replications) are known to yield variable results it was decided to make four runs. There may be variation in readings depending on which position in the machine and the particular run of the machine. Thus, we have two blocking variables: Machine Position and Run.

Run	Position 1	2	3	4	RunMean
1	31(G3)	43(G4)	67(G1)	36(G2)	44.25
2	39(G4)	96(G1)	40(G2)	48(G3)	55.75
3	57(G2)	33(G3)	40(G4)	84(G1)	53.50
4	85(G1)	46(G2)	48(G3)	50(G4)	57.25
PosMean	53	54.5	48.75	54.5	52.6875
GradeMean	G1-83	G2-44.75	G3-40	G4-43	

Source of Variation	Degrees of Freedom
Run	3 (4-1)
Position	3 (4-1)
Grade	3 (4-1)
Error	6 (15-9)
Total	15 (16-1)

An experiment in which we have t unstructured treatments and t^2 EU's which are structured depending on the values of two blocking variables is called a *basic* Latin Square Design. The square consists of

1. Rows: Levels of First Blocking Variable

2. Columns: Levels of Second Blocking Variable

There is potentially an additive effect due to the level of the row blocking variable and an additive effect due to the column blocking variable. Therefore, we need to design the experiment such that every treatment appears in every row and every column. In the example, given above note the each treatment appears in every row and column.

The randomization in a Latin Square Design is done as follows:

1. Randomly assign the numbers $1, 2, \ldots, t$ to the t Treatments.

2. Randomly assign the numbers $1, 2, \ldots, t$ to the t levels of the Row Blocking Variable.

3. Randomly assign the numbers $1, 2, \ldots, t$ to the t levels of the Column Blocking Variable.

4. Randomly assign the t^2 EU's to the t^2 positions in the $t \times t$ square.

5. Randomly select an arrangement of the Treatments within the square.

6. Randomly assign the order in which the measurements are observed.

A selection of Latin Squares from pages 307-309 in the textbook are given on the next page.

8A Appendix: Selected Latin Squares

4 × 4

```
A B C D      A B C D      A B C D      A B C D
B A D C      B C D A      B D A C      B A D C
C D B A      C D A B      C A D B      C D A B
D C A B      D A B C      D C B A      D C B A
```

5 × 5

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

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

6 × 6

```
A B C D E F      A B C D E F
B F D C A E      B A F E C D
C D E F B A      C F B A D E
D A F E C B      D C E B F A
E C A B F D      E D A F B C
F E B A D C      F E D C A B

A B C D E F      A B C D E F
B C F A D E      B A E C F D
C F B E A D      C F B A D E
D E A B F C      D E F B C A
E A D F C B      E D A F B C
F D E C B A      F C D E A B
```

27

Some Advantages of Using a Latin Square Design (LSD):

1. When the EU's are heterogeneous due to two identifiable sources of variation, a more efficient and accurate analysis can be obtained using a LSD compared to a CRD or RCBD.

2. There is greater sensitivity in the F-test for treatment effect in a LSD due to removing the sum of squares SSR and SSC from SSE. However, the reduction in SSE must be large enough to compensate for the corresponding reduction in the degrees of freedom for MSE.

3. A straight-forward analysis is available.

4. The LSD is fairly easy to implement.

5. It is possible to combine LSD's of the same size. This is very important when t is small which results in a small value for DF_{MSE}.

There are a few Disadvantages in using a LSD:

1. The number of levels of the two blocking variables must equal the number of treatments.

2. When t is small, the degrees of freedom for MSE are small and hence the power of the F-test is relatively low.

3. When there is non-additive effects of the row and column factors, that is, the row and column factors interact with the treatment effect, the LSD is not appropriate for evaluating treatment effects.

Extensions to the Standard Latin Square Design

The nature of the replication in any experimental design is crucial but this is especially important in RCBD and LSD. Consider the following example of a Latin Square Design:

Suppose four drivers and four cars are used in a study of possible differences between four gasoline additives. The four cars are identical models but there may still be slight systematic differences in their performance. Similarly, even if each driver is professionally trained and does his/her best to drive the car in the manner prescribed by the researchers, there will be systematic differences from driver to driver. The Latin square design allows for the minimization of both the car-to-car and driver-to-driver differences. In order to increase the number of data values, the design could be repeated in the following ways where a Replicate is a new running of the experiment.

1. Use the same drivers and cars in each Replicate

2. Use the same drivers but different cars in each Replicate or use the same cars but different drivers in each Replicate

3. Use different drivers and different cars in each Replicate

Case 1: Latin Square Replicated by Using Same Blocking Variables in Each Replicate:

The same four drivers and same four cars are used in the three replications of the study.

SQUARE 1						SQUARE 2						SQUARE 3				
	CARS						CARS						CARS			
DRV.	C1	C2	C3	C4		DRV.	C1	C2	C3	C4		DRV.	C1	C2	C3	C4
D1	A2	A1	A3	A4		D1	A1	A4	A3	A2		D1	A4	A2	A3	A1
D2	A1	A2	A4	A3		D2	A2	A1	A4	A3		D2	A3	A4	A1	A2
D3	A4	A3	A2	A1		D3	A3	A2	A1	A4		D3	A2	A1	A4	A3
D4	A3	A4	A1	A2		D4	A4	A3	A2	A1		D4	A1	A3	A2	A4

Source of Variation	D.F.
Square	2
Additive	3
Driver	3
Car	3
Error	36 = 47-11
Total	47

Case 2: Latin Square Replicated by Introducing Additional Versions of One of the Two Blocking Variables in the Replicates:

The same four drivers but four different cars are used in the three replications of the study.

SQUARE 1						SQUARE 2						SQUARE 3				
	CARS						**CARS**						**CARS**			
DRV.	C1	C2	C3	C4		**DRV.**	C5	C6	C7	C8		**DRV.**	C9	C10	C11	C12
D1	A3	A2	A1	A4		D1	A4	A1	A3	A2		D1	A1	A2	A3	A4
D2	A4	A1	A2	A3		D2	A1	A2	A4	A3		D2	A2	A4	A1	A3
D3	A2	A4	A3	A1		D3	A2	A3	A1	A4		D3	A3	A1	A4	A2
D4	A1	A3	A4	A2		D4	A3	A4	A2	A1		D4	A4	A3	A2	A1

Source of Variation	D.F.
Square	2
Additive	3
Driver	3
Car(Square)	9= (4-1)(3)
Error	30=47-17
Total	47

Case 3: Latin Square Replicated by Introducing Additional Versions of the two Blocking Variables in Each Replicate:

There are four new drivers and four new cars in each of the three replications of the study.

SQUARE 1						SQUARE 2						SQUARE 3				
	CARS						**CARS**						**CARS**			
DRV.	C1	C2	C3	C4		**DRV.**	C5	C6	C7	C8		**DRV.**	C9	C10	C11	C12
D1	A2	A1	A3	A4		D5	A3	A4	A1	A2		D9	A4	A2	A3	A1
D2	A1	A2	A4	A3		D6	A4	A1	A2	A3		D10	A3	A4	A1	A2
D3	A4	A3	A2	A1		D7	A1	A2	A3	A4		D11	A2	A1	A4	A3
D4	A3	A4	A1	A2		D8	A2	A3	A4	A1		D12	A1	A3	A2	A4

Source of Variation	D.F.
Square	2
Additive	3
Driver(Square)	9= (4-1)(3)
Car(Square)	9= (4-1)(3)
Error	24=47-23
Total	47

Balanced Incomplete Block Design (BIBD)

Suppose we have a set of t treatments which we want to investigate in a situation where the EU's are not homogeneous. We therefore need to use a Blocked design but only blocks containing $k < t$ EU's are available. Thus, an incomplete block design is necessary. This will result in a situation where the effect of the individual blocks is not evenly distributed over the treatments.

A degree of balance is achieved by ensuring that every pair of treatments appear together in a block equally often, but not necessarily in all blocks. The type of balance achieved is that the comparison between pairs of treatment means has the same level of precision for all possible pairs.

A **balanced incomplete block (BIB) design** is an experiment where t treatments are assigned to b blocks such that

1. Each block contains $k < t$ treatments

2. Each treatment appears in exactly r blocks

3. Every pair of treatments occurs together in exactly λ blocks

    That is every pair of treatment appears together equally often in the same block.

Thus, only a subset of the t treatments appear in every block. The number of blocks necessary to achieve balance must be a multiple of $\binom{t}{k}$, the number of ways to select k treatments to appear together in a given block. If $b = \binom{t}{k}$, then the number of blocks in which a given pair appear together is $\lambda = \binom{t-2}{k-2}$. In many cases, $\binom{t}{k}$ is much larger than the resources available to conduct the experiment. A degree of balance can be achieved with fewer than $\binom{t}{k}$ blocks. In these situations, only a subset of all $\binom{t}{k}$ possible pairings are selected but the number of blocks in which each pair of treatments appear together remains a constant $\lambda < \binom{t-2}{k-2}$.

We will use the following notation in a BIBD:

1. t = number of treatments (factor level combinations)

2. b = number of blocks

3. r = number or EU's/trt, that is, number of blocks in which each treatment appears

4. k = number of EU's per block

5. λ = number of blocks in which each pair of treatments appear together

6. N = number of EU's in the experiment

There are a number of restrictions that must be satisfied in order for a BIBD to exist:

R1. $N = tr = bk$

- number of EU's is (# Trts)*(# Reps) or (# Blocks)*(# EU's/Block)

R2. $\lambda = \frac{r(k-1)}{t-1}$ with λ an integer

- Each Trt is paired in λ blocks with each of the other $t-1$ trts

- Each Trt appears in $\lambda(t-1)$ pairs of Trts

- Each Trt is paired with $k-1$ Trts in r blocks, that is, each Trt appears in r blocks with k Trts per block

- Each treatment appears in $r(k-1)$ pairs of Trts

- Therefore, $\lambda(t-1) = r(k-1)$

R3. $\lambda < r < b$

- $k < t \implies \lambda = \frac{r(k-1)}{t-1} < r$ and $r < b$ because no treatment appears in every block

R4. $b \geq t$

- See Hinkelmann and Kempthorne (2005), *Design and Analysis of Experiments, Vol. 2.*

Consider the following examples.

Example 1: Suppose we have the following arrangement of $t = 6$ treatments $(T_1, T_2, T_3, T_4, T_5, T_6)$ in $b = 4$ blocks:

Blocks

B_1	B_2	B_3	B_4
T_1	T_1	T_2	T_4
T_2	T_3	T_3	T_5
T_4	T_6	T_5	T_6

Is this a BIB design?

Example 2: Suppose we have the following arrangement of $t = 6$ treatments $(T_1, T_2, T_3, T_4, T_5, T_6)$ in $b = 10$ blocks:

Blocks

B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9	B_{10}
T_1	T_1	T_2	T_4	T_1	T_1	T_1	T_2	T_2	T_3
T_2	T_3	T_3	T_5	T_2	T_3	T_4	T_3	T_5	T_4
T_4	T_6	T_5	T_6	T_6	T_5	T_5	T_4	T_6	T_6

Is this a BIB design?

Weight Gain of Rabbits Under Six Diets

A study of the difference of 6 proposed Diets on the weight gain of young rabbits is proposed. Because weight varies considerably amongest young rabbits, it is proposed to block the experiment based on litters. There are 10 litters of rabbits available of varying sizes. The minimum litter size is 3. Therefore, only 3 of the 6 diets can be observed in any particular litter. Is a BIBD possible in this situation?

$$t = 6, \quad k = 3, \quad b = 10, \quad r = 5 \Rightarrow N = tr = 30 = bk, \quad \lambda = \frac{r(k-1)}{t-1} = \frac{5(3-1)}{6-1} = 2$$

Thus, a BIBD with 5 reps of the 6 diets can be run with every pair of diets observed together is 2 litters. In order to have complete balance we would need a multiple of $\binom{6}{3} = 20$ litters. In this experiment, we have only 10 litters thus every triple of diets will not be observed. The actual experiment yielded the following results:

Litter	Diet 1	2	3	4	5	6	Litter Totals	Litter Means
1		32.6	35.2			42.2	110.0	36.67
2	40.1	38.1	40.9				119.1	39.70
3			34.6	37.5		34.3	106.4	35.47
4	44.9		43.9		40.8		129.6	43.20
5			40.9	37.3	32.0		110.2	36.73
6		37.3			40.5	42.8	120.6	40.20
7	45.2	40.6		37.9			123.7	41.23
8	44.0				38.5	51.9	134.4	44.80
9		30.6		27.5	20.6		78.7	26.23
10	37.3			42.3		41.7	121.3	40.43
Diet Totals	211.5	179.2	195.5	182.5	172.4	212.9	1154.0	
Diet Means	42.3	35.84	39.1	36.5	34.48	42.58		38.47

Source of Variation	D.F.
Litter	9
Diet	5
Error	15 = 29-14
Total	29

I. Split Plot Designs

In some factorial experiments it may be impossible to completely randomize all the treatments to the EU's. This often results in a generalization of the factorial design called the **Split Plot Design**.

Consider a paper manufacturer who is interested in studying the effects of three different pulp preparation methods and four different cooking temperatures for the process on the quality of paper produced:

Factor 1 - Pulp Preparation Methods: M1, M2, M3

Factor 2 - Cooking Temperatures: T1=$220°F$, T2=$240°F$, T3=$260°F$, T4=$280°F$

The process chemist wants to study the effect of pulp preparation and temperature on the shear strength of the paper. Each replicate of a factorial experiment requires 12 observations, and the experimenter has decided to run three replicates, thus needs a total of 36 runs of the process. However, the pilot plant is only capable of making 12 runs per day, so the experimenter decides to run one replicate on each of the three days, hence days are blocks.

A further complication is that it is very expensive to make individual batches of the pulp by a given preparation method. Therefore, a single batch of pulp is produced by each of the three methods on each day in random order. Next, each of the three batches are divided into four sub-batches. The four sub-batches are then randomly assigned to the four temperatures and after the process is completed, the shear measurements are made. An outline of the experiment is given below the Runs labeled as R1 - R12 on each of the three Days.

	Day 1			Day 2			Day 3	
M1	M2	M3	M2	M1	M3	M3	M1	M2
R1-T1	R5-T4	R9-T3	R1-T1	R5-T4	R9-T1	R1-T3	R5-T2	R9-T3
R2-T2	R6-T3	R10-T2	R2-T4	R6-T1	R10-T2	R2-T1	R6-T3	R10-T1
R3-T3	R7-T2	R11-T4	R3-T2	R7-T2	R11-T3	R3-T4	R7-T4	R11-T2
R4-T4	R8-T1	R12-T1	R4-T3	R8-T3	R12-T4	R4-T2	R8-T1	R12-T4

Is this design might a RCBD with a 3x4 factorial treatment structure?

Blocks (3-Days)

3 levels of Preparation Method (M1, M2, M3) Crossed with

4 levels of Temperature (T1, T2, T3, T4).

In order for this to be a randomized complete block design, it would be necessary that

1. 12 batches prepared each day - four batches for each preparation method.

2. On each Day, randomly assign the 12 Treatment combinations (a Preparation Method and a Temperature) to the 12 runs of the process

3. After the process is finished, obtain a tensile strength measurement for each batch.

4. Conduct a new randomization on each of the three days.

The design would appear as follows:

Day 1			Day 2			Day 3		
R1-M2T3	R5-M3T4	R9-M2T1	R1-M3T4	R5-M2T1	R9-M1T2	R1-M3T2	R5-M1T3	R9-M3T4
R2-M3T2	R6-M1T1	R10-M1T3	R2-M2T3	R6-M3T2	R10-M2T4	R2-M1T1	R6-M2T4	R10-M2T2
R3-M3T3	R7-M2T2	R11-M2T4	R3-M1T4	R7-M1T3	R11-M1T1	R3-M1T2	R7-M3T1	R11-M3T3
R4-M1T4	R8-M1T2	R12-M3T1	R4-M3T1	R8-M3T3	R12-M2T2	R4-M2T3	R8-M2T1	R12-M1T4

However, this was not how the experiment was conducted.

The chemist made up a batch of pulp using a randomly selected method of preparation, divided the batch into four sub-batches, and then randomly assigned a temperature to the sub-batches. Because of the economics of preparing batches of pulp and the size of the batches, this was the only feasible manner to conduct the experiment.

A randomized complete block design would require 36 batches of pulp (12 per day), which is not economically feasible. The split-plot design requires only 9 batches (3 batches per day). Thus, the split-plot design resulted in a considerable cost savings in conducting the experiment.

However, there is a cost in terms of statistical efficiency with respect to the power of F-test in testing hypotheses involving main effects and interactions for the whole plot treatments. This can be seen in the AOV tables for the two designs:

Randomized Block Completely Randomized AOV

Source of Variation	DF
Days	2
Method	2
Temp	3
Method*Temp	6
Error	22 (35-13)
Total	35

Randomized Block Split-Plot AOV

Source of Variation	DF
Days	2
Method	2
Error(A) = Days*Method	4
Temp	3
Method*Temp	6
Error(B)	18 (35-17)
Total	35

The Error Terms for testing the difference in the three Methods of preparation have greatly different degrees of freedom for the two designs:

The RCBD with Method and Temperature completely randomly assigned has an Error term for testing Method differences with degrees of freedom equal to 22.

The RCBD with a Split-Plot Treatment assignment of the levels of Method and Temperature, the Error term for testing Method differences has degrees of freedom equal to 4.

The power of the F-test is considerably larger for an F-statistic having df (2, 22) df versus one having (2, 4) df.

The F-tests for the main effect of Temperature and the interaction have df equal to (3, 22) for Temp and (6, 22) for Method*Temp in the Completely randomized treatment assignment in the RCBD, whereas

The F-tests for the main effect of Temperature and the interaction have df equal to (3, 18) for Temp and (6, 18) for Method*Temp in the Split-Plot treatment assignment in the RCBD

Thus, for all three effects, main effect of Preparation Method, main effect of Temperature and the interaction between Preparation Method and Temperature, the power of the F-tests would be large in the completely randomized assignment of the treatments in comparison to the Split-plot method of assigning the treatments.

We can thus conclude that the Split Plot assignment methodology should only be used when the physical nature of the experiment requires using a Split Plot assignment.

A split-plot design is essentially two experiments superimposed on each other. One experiment has the whole plot factor applied to the larger experimental units (or it is a factor whose levels are hard or expensive to change) and the other experiment has the subplot factor applied to the smaller experimental units (or it is a factor whose levels are easier or less costly to change). That is, the Whole Plot factor's levels are assigned to the whole plots in a randomized complete block design. Each whole plot is then subdivided into b subunits and the levels of subplot factor are assigned to the subunits as in a completely randomized design.

A1	A2	A3	
B3	B1	B4	**Subplot Factor B**
B2	B4	B1	
B4	B3	B2	
B1	B2	B3	
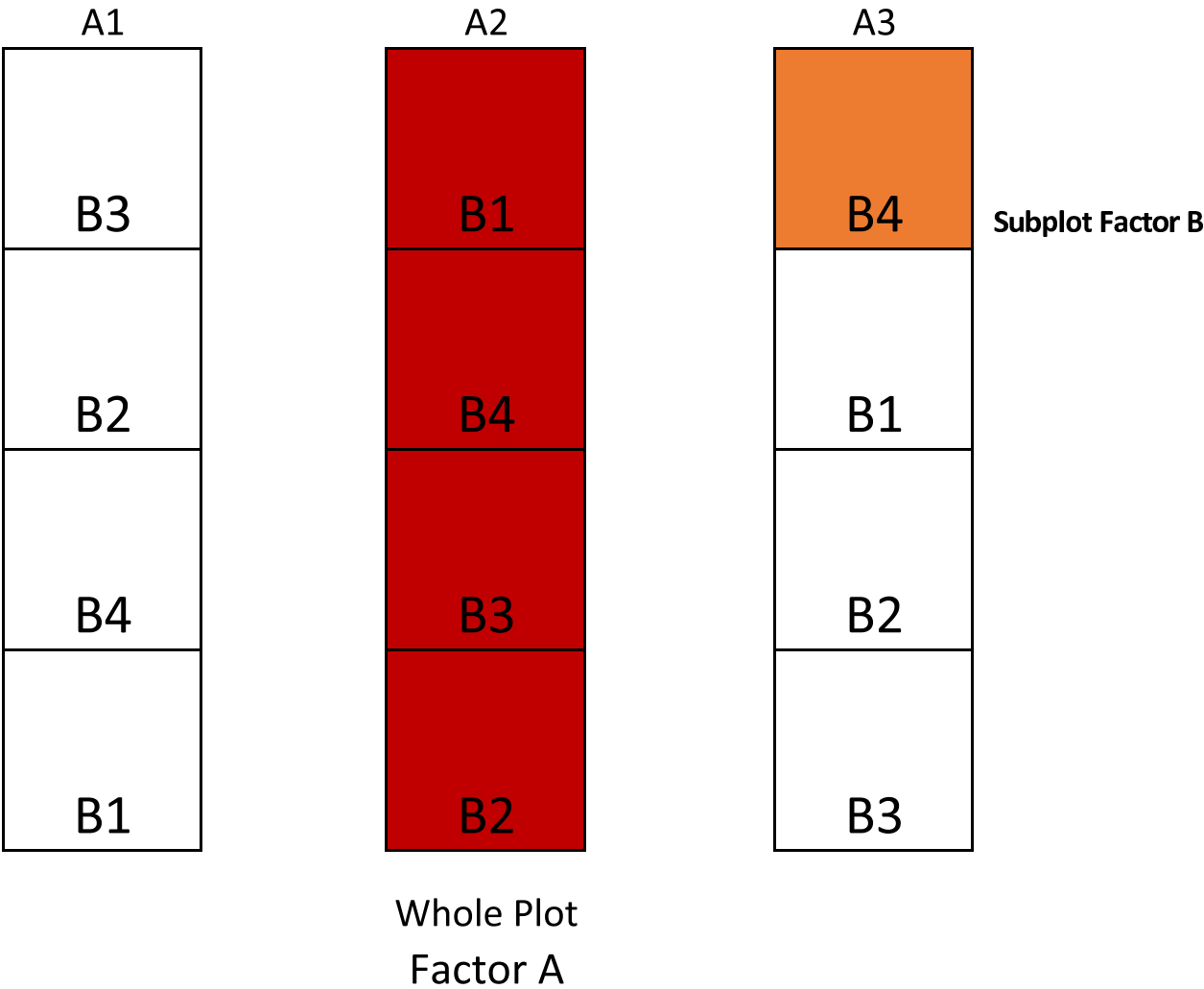
Whole Plot
Factor A

In our example,

the whole plot EU's are the batches which are randomly assigned to the three Preparation Methods

the split plot EU's are the four sub-batches obtained from each of the batches and are randomly assigned to the four Temperatures

Randomized Complete Block Split Plot Design Example

The experiment involves three Varieties of alfalfa (Ladak, Ranger, Cossack) and four different cutting dates (A=none, B=Sept. 1, C=Sept. 20, D=Oct. 7). There were 6 fields involved in the experiment which will be treated as blocks. Each field was divided into 3 plots and the varieties of alfalfa were randomly assigned to the plots. Each plot was then divided into 4 subplots and the Cutting Dates were randomly assigned to the subplots within each plot. The cutting date represents the date on which the 3rd cutting of the alfalfa field was made in the previous year.

Thus the plots of land are the EU's for Varieties and

the subplots within a plot are the EU's for the Cutting Dates.

The randomization for the plots and subplots is given below:

	Field 1			Field 2			$\cdots$	Field 6		
	PLOT1	PLOT2	PLOT3	PLOT1	PLOT2	PLOT3	$\cdots$	PLOT1	PLOT2	PLOT3
	Ladak	Cossack	Ranger	Cossack	Ladak	Ranger	$\cdots$	Cossack	Ladak	Ranger
SUB1	NONE	S1	S20	S1	NONE	07	$\cdots$	S20	O7	NONE
SUB2	07	S20	070	S20	O7	S1	$\cdots$	S1	S20	07
SUB3	S20	NONE	NONE	01	S1	S20	$\cdots$	O7	S1	S1
SUB4	S1	07	S1	NONE	S20	NONE	$\cdots$	NONE	NONE	S20

The total yield (Tons/acre) for three cuttings in the current year are given in the following table. Do the yields vary across the factors cutting date and varieties?

Variety	Date	Fields						TrT Mean
		1	2	3	4	5	6	$\bar{Y}_{ij.}$
Ladak	None	2.17	1.88	1.62	2.34	1.58	1.66	1.8750
	S1	1.58	1.26	1.22	1.59	1.25	0.94	1.3067
	S20	2.29	1.60	1.67	1.91	1.39	1.12	1.6633
	O7	2.23	2.01	1.82	2.10	1.66	1.10	1.8200
Cossack	None	2.33	2.01	1.70	1.78	1.42	1.35	1.7650
	S1	1.38	1.30	1.85	1.09	1.13	1.06	1.3017
	S20	1.86	1.70	1.81	1.54	1.67	0.88	1.5767
	O7	2.27	1.81	2.01	1.40	1.31	1.06	1.6433
Ranger	None	1.75	1.95	2.13	1.78	1.31	1.30	1.7033
	S1	1.52	1.47	1.80	1.37	1.01	1.31	1.4133
	S20	1.55	1.61	1.82	1.56	1.23	1.13	1.4833
	O7	1.56	1.72	1.99	1.55	1.51	1.33	1.6100

AOV Table

Source of Variation	DF
Field	5
Variety	2
Error(A) = Field*Variety	10
Date	3
Variety*Date	6
Error(B)	45 (71-26)
Total	71

Completely Randomized Split-Plot Design Example

The superintendent of the city schools in a large midwestern city decided to investigate the effect of three teaching methods on young student achievement in mathematics concepts:

1. Classical classroom instruction

2. Web-based instruction

3. Classical classroom instruction plus Web-based instruction

He was also concerned about the effect of the use of calculators on achievement in mathematics concepts. Because the teaching methods must naturally be applied to a block of students, the experiment was designed in the following manner. Twelve third grade classrooms were randomly selected with four classrooms randomly assigned to each of the three teaching methods. Ten students in each classroom were randomly selected and given calculators and ten students were randomly selected and were not allowed to use calculators. Prior to receiving instruction and at the end of the instruction period, each student was given a test and the students gain in mathematical achievement was recorded. This is a complete randomized split plot design with Whole Plot being the classroom and the Split Plot being the individual students. The Whole Plot treatment is method of instruction and the Split Plot treatment is calculator use. Note, this study could have been designed as a randomized block split plot design by randomly selecting four elementary schools throughout the city, then randomly selecting three fourth grade classes within each school with the Method of Instuction randomly assigned within each school.

How would the randomization differ in a Completely Randomized Split-Plot Design from the randomization in a Randomized Complete Block Split-Plot design?

Randomized Complete Block Split-Split-Plot Design

When we have three factors, Factor A assigned to Whole Plot EU's, Factor B assigned to Split-Plot EU's, Factor C assigned to Split-Split-Plot EU's, a subdivision of the subplots is required with all levels of Factor C randomly assigned to these new subdivisions, referred to as sub-subplots. There are r blocks containing r Whole Plot EU's each. The design is referred to as a **Randomized Complete Block Split-Split-Plot Design** and the design has three different sizes of EU's.

EXAMPLE   In an industrial experiment, large batches of four alloys (A1, A2, A3, A4) are prepared in a furnace. After a batch is removed from the furnace, it is subdivided into small containers for application at different cooling temperatures (T1, T2, T3). After each container of alloy is cooled to its proper temperature, the alloy is further subdivided for application of one of five levels of an anti-oxidizing agent (O1, O2, O3, O4, O5). Identify the factors and relevant EU's in this study. Describe the appropriate randomization procedures.

Strip-Plot Design

Another variation on the Split-Plot design occurs when the subunit treatments occur in a strip across the wholeplot units. The Strip-plot design can be useful in agricultural field studies when two treatment factors can only be applied across large field plots. The levels of Factor A are randomly assigned to the plots in a randomized complete block design. The plots for Factor B are constructed in the same manner but are laid out perpendicular to the plots for Factor A. The levels of Factor B are then randomly assigned to this second array of plots across the same block. A diagram displaying this arrangement is given below:

	A2	A1	A3	EU for Factor A
B3				
B1				
B2				
B4				

EU for Factor B

EU for
A*B Interaction
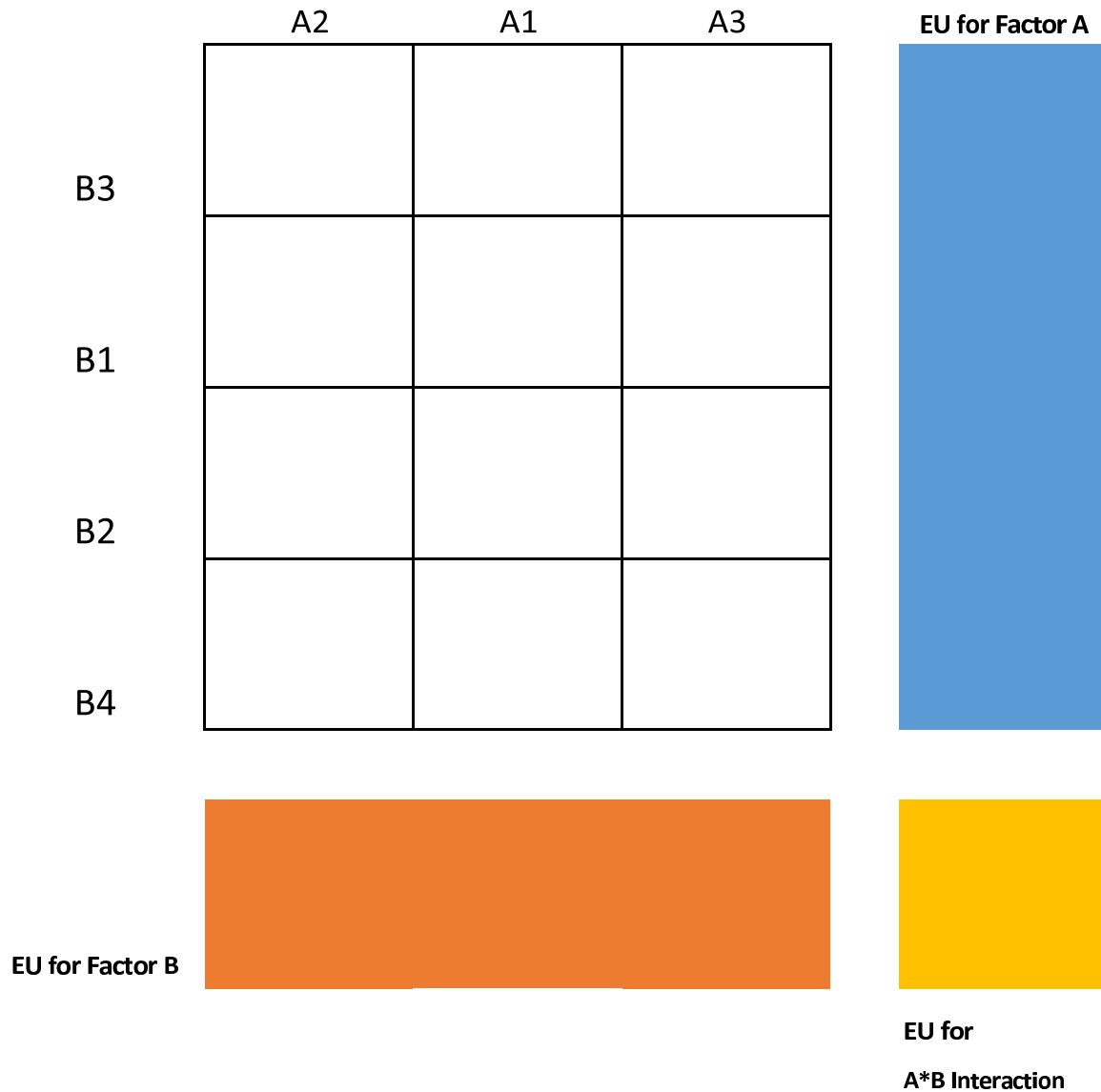
The Strip-Plot design has three sizes for experimental units where the units for the main effects of Factors A and B are equivalent to whole plots, but each with a different orientation. The EU's for the A*B interaction effect is a subplot where there is an intersection of the two whole plots for the respective levels of Factors A and B.

Questions: How would the randomization differ in the Alfalfa example so that the experiment would be a Strip Plot Design instead of a Split Plot design?

Alfalfa experiment as a **Split Plot Design:**

WP-Factor is Variety with EU=Plot and

SP-Factor is Date with EU=SUB

	Field 1			Field 2			$\cdots$	Field 6		
	PLOT1	PLOT2	PLOT3	PLOT1	PLOT2	PLOT3	$\cdots$	PLOT1	PLOT2	PLOT3
	Ladak	Cossack	Ranger	Cossack	Ladak	Ranger	$\cdots$	Cossack	Ladak	Ranger
SUB1	NONE	S1	S20	S1	NONE	07	$\cdots$	S20	O7	NONE
SUB2	07	S20	070	S20	O7	S1	$\cdots$	S1	S20	07
SUB3	S20	NONE	NONE	01	S1	S20	$\cdots$	O7	S1	S1
SUB4	S1	07	S1	NONE	S20	NONE	$\cdots$	NONE	NONE	S20

Alfalfa experiment as a **Strip Plot Design:**

Factor A is Variety with levels randomly assigned to EU's being the Plots running North to South and

Factor B is Date with levels randomly assigned to EU's being the Plots running East to West

	Field 1			Field 2			$\cdots$	Field 6		
	PLOT1	PLOT2	PLOT3	PLOT1	PLOT2	PLOT3	$\cdots$	PLOT1	PLOT2	PLOT3
	Ladak	Cossack	Ranger	Cossack	Ladak	Ranger	$\cdots$	Cossack	Ladak	Ranger
PLOT1	O7	07	O7	S20	S20	S20	$\cdots$	O7	O7	O7
PLOT2	S20	S20	S20	NONE	NONE	NONE	$\cdots$	S1	S1	S1
PLOT3	S1	S1	S1	O7	O7	O7	$\cdots$	NONE	NONE	NONE
PLOT4	NONE	NONE	NONE	S1	S1	S1	$\cdots$	S20	S20	S20

EXAMPLE An engineer wants to investigate the strength of ceramic components made with three different percentages of silicon (P1, P2, P3) and four heating temperatures (T1, T2, T3, T4). The process is to repeated on five consecutive days. The engineer has only one kiln for heating the ceramic material and the material must be heated for five hours. Therefore, she can conduct at most four heat treatments per day. On each day, a single batch of the ceramic material is made with each of the silicon percentages. The three batches are subdivided into four sub-batches and one sub-batch from each of the three silicon percentages is then placed in the kiln at one of the four temperatures. The four ceramic sub-batches are then heated for 5 hours. Thus, on a given day, three batches of ceramic material are made and the kiln is run for five hours at each of the four temperatures.

- Is this a RCBD with a 3x4 treatment structure?

- Is this a RCB split-plot Design with a 3x4 treatment structure?

- Is this a RCB split-split-plot Design with a 3x4 treatment structure?

- Is this a RCB strip-plot Design with a 3x4 treatment structure?

How to Identify a Split-Plot Design

There many other variation on the basic split-plot design. The crucial elements to consider are

1. What restrictions have been placed on the randomization

2. Are there different size EU's for the different factors

3. To which EU's were the levels of each factor randomly assigned.

4. Is there true replication or do we have subsampling within the basic design.

Repeated Measures Designs

Repeated Measures designs are experiments or studies in which the experimental unit, EU's, have several measurements taken on them either spatially or temporally. This is in contrast to the standard design in which a single measurement is taken on the experimental unit. Several examples of this type of study are subsampling or crossover designs. In a study involving subsampling there is not generally any other variable of interest associated with the repeated measurements. The goal is to evaluate the measuring device or to evaluate the variability across the surface of an EU, for example, a field of corn or silicon wafer. In the crossover design the repeated measures occur because the EU actually is observed or measured after each application of a **NEW** treatment. A further application of repeated measures is when an EU receives a treatment and the experimenter measures a response at specific time points or a specific locations on the EU. In this type of situation, the variable Time or Location is of interest to the experimenter. By using the Repeated Measures Analysis the researcher will be able to determine time trends or spatial patterns and determine if there certain trends or patterns vary amongst the different treatments. For instance, a drug company may be interested in knowing if their drug reduces blood pressure linearly with respect to time t and if after time t, a person's blood pressure remains constant. A wafer manufacturer may want to know if there is a different pattern in the conductivity of the wafer across the wafer depending of the type of protective coating placed on the wafer.

The following example from Crowder and Hand's *Repeated Measures* book will be used to illustrate this design. In their study, three levels of a vitamin E supplement: Zero (Control), Low, High were given to guinea pigs. Five pigs were randomly assigned to each of the three levels of Vitamin E supplement. The weight of the pigs were recorded at 1, 2, 3, 4, 5, 6 weeks after the beginning of the study. This is a repeated measures experiment because each pig, the EU, is given only one treatment but each pig is measured 6 times. The experimenter is interested in the trend in weight over time.

Weight of Guinea Pigs Under 3 Levels of Vitamin E

Level of E	Animal	Week1	Week2	Week3	Week4	Week5	Week6
C	1	455	460	510	504	436	466
C	2	467	565	610	596	542	587
C	3	445	530	580	597	582	619
C	4	485	542	594	583	611	612
C	5	480	500	550	528	562	576
L	6	514	560	565	524	552	597
L	7	440	480	536	484	567	569
L	8	495	570	569	585	576	677
L	9	520	590	610	637	671	702
L	10	503	555	591	605	649	675
H	11	496	560	622	622	632	670
H	12	498	540	589	557	568	609
H	13	478	510	568	555	576	605
H	14	545	565	580	601	633	649
H	15	472	498	540	524	532	583

Note that there are three treatments with $r = 5$ reps per treatment for a total of 15 EU's (pigs) each of which is weighed 6 times for a total of 90 observations. In contrast, a completely randomized design with 90 observations would have 90 EU's each weighed once. Thus, 75 more pigs are required to perform the CRD. However, this gain in economy has limitations. The inferences are being made to a population of pigs. In the repeated measures design only 15 pigs from the population are being observed. Thus, there may be greater variability in the estimation of the treatment means due to having such a small sample size per treatment. On the other hand the repeated measures design allows the researcher to tract the behavior of the individual pig over the 6 weeks and hence provides information concerning the potential differences in fluctuations in weight for the individual pigs.

A CROSS-OVER DESIGN

In a Repeated Measures Experiment, the EU receives a Treatment and then the EU has multiple observations or measurements made on it over time or space. The EU **Does Not** receive a New Treatment between Measurements.

In a Crossover Design, each EU is observed under each of the t Treatments during t observation times. That is, every EU has multiple treatments applied to it and then a new measurement or observation is obtained.

Because the Treatments are compared on the Same EU's, the between EU variation is greatly reduced. The individual EU's serve as blocks in order to reduce the experimental variation (reduced SSE) and hence increase the efficiency of the estimation of the treatment means.

When comparing treatments, the effect of the Time Period in which the treatment was applied comes into the analysis. Differences in observations may be due to treatment differences and/or time period differences. CrossOver Designs are constructed to avoid confounding the Time Period Effects with the Treatment Effects:

Suppose we have 3 treatments: T_1, T_2, T_3 with each treatment applied to each of 12 patients during 3 time periods: P_1, P_2, P_3. The drugs were applied in the same order to all 12 patients:

Patient	Time Period P1	P2	P3
1	T_1	T_2	T_3
2	T_1	T_2	T_3
3	T_1	T_2	T_3
4	T_1	T_2	T_3
5	T_1	T_2	T_3
6	T_1	T_2	T_3
7	T_1	T_2	T_3
8	T_1	T_2	T_3
9	T_1	T_2	T_3
10	T_1	T_2	T_3
11	T_1	T_2	T_3
12	T_1	T_2	T_3

From the data, a large difference was observed in the treatment means: $\bar{y}_{1..}, \bar{y}_{2..}, \bar{y}_{3..}$. Is this difference due to Treatment Differences or Time Period Differences? With the above design, it would be impossible to determine. The sample mean responses for evaluation the differences in the effect of the 3 Time Periods are identical to the 3 Treatment Means. That is, with this design, the effects of Treatment and Time Period are Confounded.

Thus, it is necessary to consider multiple Sequences in which the Treatments are administered to the EU's: There are $3! = 6$ possible sequences in which the three treatments could be administered to the 12 subjects during the three treatment periods.

	Time Period		
Sequence	P1	P2	P3
S1	T_1	T_2	T_3
S2	T_2	T_3	T_1
S3	T_3	T_1	T_2
S4	T_2	T_1	T_3
S5	T_3	T_2	T_1
S6	T_1	T_3	T_2

The experimenter could then randomly assign 2 patients to each of the 6 sequences. Then, there would not be confounding between the effects due to treatments, sequences, and time period. Every treatment is observed in every sequences and in every time period.

EXAMPLE: Duration Effect of Three Formulations of a Drug

Twelve males volunteered to participate in a study to compare the effect of three formulations of a drug product: Formulation 1 was a 5-mg tablet, Formulation 2 was a 100-mg tablet, and Formulation 3 was a sustained-release capsule.

The experimenter selected the first 3 of the 6 sequences and randomly assigned 4 subjects to each sequence.

On each treatment day, volunteers were given their assigned formulation and were observed to determine the duration of effect (blood pressure lowering). The experimental data is given here.

Sequence	Patient(Seq.)	Time Period P1	P2	P3	PA(S) $\bar{Y}_{ij.}$	Sequence $\bar{Y}_{i..}$
	1	2.5	3.2	4.4	$3.3\bar{6}$	
S1	2	2.0	2.6	3.1	$2.5\bar{6}$	$2.38\bar{3}$
$T_1 \to T_2 \to T_3$	3	1.6	2.7	3.2	2.5	
	4	0.1	1.3	1.9	1.1	
	5	2.5	3.5	1.9	$2.6\bar{3}$	
S2	6	3.8	4.1	2.5	$3.4\bar{6}$	2.55
$T_2 \to T_3 \to T_1$	7	2.7	2.9	2.4	$2.6\bar{6}$	
	8	1.4	1.6	1.3	$1.4\bar{3}$	
	9	3.3	1.9	2.7	$2.6\bar{3}$	
S3	10	2.1	0.6	1.5	1.40	$2.558\bar{3}$
$T_3 \to T_1 \to T_2$	11	4.6	3.3	3.2	3.70	
	12	3.0	2.5	2.0	2.50	

C3: The Measurement Process

The measurement process is the method by which the response is recorded on the EU after the EU has been randomly to a treatment

1. Single measurement on each of the experimental units

1. Repeated measurements on experimental unit over time (Longitudinal Measurements): EU (patient or lab animal) is randomly assigned one of 6 possible dose levels (treatment) of a drug . The amount of drug in the blood stream is measured every hour over the next 2 days. A total of 48 measurements per EU.

2. Repeated measurements at various locations on experimental unit (Spatial Measurements): Studying three electroplating processes(3 treatments). Six specimens of metal (EU), all uniform in size and thickness, are randomly assigned to each of the 3 electroplating processes (treatments). After electroplating, the thickness of each specimen was measured at three specified locations on the specimen.

3. Sub sampling of experimental unit: In some experiments the researcher wants to obtain additional information concerning the variability of the response within the individual EU's. The EU may be subdivided and/or random samples of information are obtained on the each EU. An experiment was designed to study amount of residual pesticide remaining at harvest on lettuce plants. Four 1 acre plots are randomly assigned to receive one of five different types of pesticides. Twenty lettuce plants were randomly selected from each plot and the amount of pesticide residual was measured on each plant.

Examples

For each of the following experiments, we will specify the following components:

1. Type of Randomization, for example, CRD, RCBD, LSD, BIBD, SPLIT-PLOT, Crossover, etc.;

2. Type of Treatment Structure, for example, Single Factor, Crossed, Nested, Fractional, etc.;

3. Identify each of the factors as being Fixed or Random;

4. Describe the Experimental Units and Measurement Units.

5. Describe the Measurement Process: Response Variable, Covariates, SubSampling, Repeated Measures

Experiment 1: A laboratory study of the stress (psi) of titanium is to be designed involving laboratories in the United States and Germany. Three laboratories are randomly selected from the many laboratories within each of the two countries. Two temperatures($100, 200^0F$), and four strain rates (1, 10, 100, 1000 $\sec^{-1}$) are to be investigated. Two titanium specimens are randomly assigned to each lab-temperature-strain combination and a stress reading is made on each specimen.

Experiment 2: A textile specialist is interested in the effect of Breed of Sheep, Season of the year, and Ranch on the production of wool. She decides to investigate the variability in length of wool fibers from 4 breeds of sheep during each of the 2 harvesting seasons. For each breed, 8 ranches were randomly selected from a listing of ranches raising that breed of sheep. During each of the 2 harvesting seasons, a random sample of 5 sheep was selected at each of these ranches. The age of each of each sheep at the beginning of the study was recorded since the sheep ranged in age from 2-15 years. On each selected sheep, the wool length was determined at 4 randomly selected sites.

Experiment 3: A biologist conducted an experiment to investigate changes in the plankton food change of a pond when a photosynthesis-inhibiting herbicide was introduced into the pond. Water was obtained from three ponds. From each of the three ponds, four containers holding 8 gallons of well-mixed pond water were obtained. The four containers were randomly assigned to be dosed with one of the following rates of herbicide: 0, 0.1, 0.5, 1.0 mg/liter. The large containers were divided into eight 1-gallon glass jugs filled with the well-mixed, dosed pond water, sealed and suspended in the pond just below the water surface. Bottles were given labels so that two bottles of each dose could be removed at the end of the first day, at the end of first week, second week, and finally third week. Rotifers are a major element of the plankton food change in this pond. The number of rotifers present in the bottle were counted immediately after the bottle was removed from the pond. The counts in numbers of rotifers/liter are given here with the following notation: Pond (P), Week (W), Dose (D), and Bottle (B).

Dose	Bottle	Pond 1 Week				Pond 2 Week				Pond 3 Week			
		0	1	2	3	0	1	2	3	0	1	2	3
0	1	6718	5166	3815	2340	6222	4656	3351	1804	7081	5466	4151	2604
	2	6392	5039	3382	2881	5891	4527	2828	2318	6629	5393	3628	3118
0.1	1	5832	5233	4373	2453	5323	4722	3837	1935	6123	5533	4637	2735
	2	5569	4350	3333	2924	5071	3849	2833	2442	5896	4605	3633	3242
0.5	1	5924	3953	4912	2575	5423	3435	4421	2057	6242	4235	5221	2857
	2	6715	4295	4427	4211	6205	3759	3972	3711	7051	4559	4772	4511
1.0	1	6821	4849	4900	4451	6316	4394	4400	3915	7112	5194	5200	4715
	2	5871	4763	4067	4552	5368	4236	3576	4025	6117	5036	4376	4825

Experiment 4: The following data is from Gennings, Chinchilli, and Carter (1989) *Journal of the American Statistical Association*, Vol. 84, pp. 805-809. An in vitro toxicity study of isolated hepatocyte suspensions was conducted to study the impact of combining carbon tetrachloride (CCl_4) and chloroform ($CHCl_3$) on the toxicity of cells. Cell toxicity was measured by the amount of lactic dehdrogenase (LDH) enzyme leakage. The study involved randomly assigning four flasks to each of the 16 treatments obtained by combining four levels of CCl_4: 0, 1.0, 2.5, 5.0 mM with four levels of $CHCl_3$: 0, 5, 10, 25 mM. The percent LDH leakage from the cells in each of the 64 flasks was measured just prior to applying the treatment to the flasks and at .01, .25, .5, 1, 2, and 3 hours after applying the treatment. The percent LDH leakage is given in the following table.

CCl_4	$CHCl_3$	Time since treatment (Hours)							CCl_4	$CHCl_3$	Time since treatment (Hours)						
		0	.01	.25	.5	1.0	2.0	3.0			0	.01	.25	.5	1.0	2.0	3.0
0	0	.08	.09	.09	.08	.10	.10	.12	0	0	.07	.08	.08	.08	.09	.09	.10
0	0	.08	.10	.10	.09	.12	.15	.13	0	0	.06	.08	.06	.07	.08	.10	.11
0	5	.06	.11	.14	.12	.14	.13	.12	0	5	.05	.07	.13	.08	.10	.10	.12
0	5	.11	.14	.16	.18	.20	.21	.14	0	5	.06	.06	.07	.13	.14	.15	.16
0	10	.06	.11	.20	.36	.46	.44	.46	0	10	.06	.07	.17	.18	.21	.22	.22
0	10	.08	.14	.24	.27	.29	.32	.34	0	10	.05	.05	.15	.16	.19	.22	.23
0	25	.07	.10	.25	.51	.65	.66	.70	0	25	.07	.07	.17	.24	.34	.37	.41
0	25	.11	.11	.33	.39	.48	.52	.55	0	25	.07	.06	.16	.24	.31	.36	.41
1	0	.06	.11	.13	.09	.10	.11	.11	1	0	.05	.08	.10	.10	.11	.12	.13
1	0	.08	.14	.15	.14	.16	.19	.21	1	0	.05	.09	.08	.09	.11	.12	.13
1	5	.05	.13	.18	.37	.41	.42	.46	1	5	.06	.10	.14	.16	.16	.20	.18
1	5	.10	.16	.22	.22	.29	.30	.21	1	5	.05	.08	.15	.18	.19	.21	.21
1	10	.06	.10	.25	.61	.57	.60	.63	1	10	.05	.07	.24	.27	.29	.32	.32
1	10	.11	.14	.26	.30	.30	.35	.29	1	10	.05	.06	.16	.21	.24	.27	.27
1	25	.07	.09	.23	.39	.58	.53	.67	1	25	.06	.06	.15	.22	.30	.44	.56
1	25	.08	.11	.28	.40	.42	.75	.72	1	25	.06	.05	.15	.27	.36	.43	.55
2.5	0	.06	.09	.19	.56	.64	.33	.34	2.5	0	.05	.08	.18	.19	.19	.21	.20
2.5	0	.10	.10	.19	.21	.23	.28	.23	2.5	0	.05	.10	.21	.23	.28	.29	.31
2.5	5	.07	.10	.22	.57	.62	.66	.70	2.5	5	.06	.08	.19	.23	.24	.27	.31
2.5	5	.12	.11	.24	.28	.30	.35	.30	2.5	5	.06	.07	.21	.25	.28	.30	.32
2.5	10	.05	.12	.28	.33	.43	.49	.58	2.5	10	.06	.09	.33	.26	.31	.34	.36
2.5	10	.08	.14	.23	.37	.43	.47	.40	2.5	10	.06	.09	.19	.23	.29	.34	.34
2.5	25	.05	.07	.22	.59	.65	.67	.67	2.5	25	.04	.05	.21	.29	.36	.54	.72
2.5	25	.09	.09	.24	.31	.35	.46	.45	2.5	25	.05	.04	.15	.25	.36	.40	.48
5	0	.06	.09	.52	.77	.78	.73	.76	5	0	.06	.08	.45	.50	.49	.60	.71
5	0	.08	.09	.60	.60	.57	.73	.79	5	0	.06	.10	.42	.44	.62	.62	.73
5	5	.05	.11	.21	.27	.30	.36	.41	5	5	.05	.10	.20	.22	.24	.28	.33
5	5	.09	.12	.21	.22	.27	.32	.28	5	5	.05	.08	.17	.21	.26	.27	.32
5	10	.04	.10	.24	.26	.33	.39	.47	5	10	.06	.09	.25	.29	.33	.37	.40
5	10	.11	.11	.23	.27	.31	.36	.31	5	10	.05	.05	.12	.16	.22	.27	.29
5	25	.07	.07	.21	.55	.60	.66	.66	5	25	.05	.05	.23	.31	.35	.53	.66
5	25	.08	.09	.23	.31	.41	.58	.67	5	25	.06	.04	.12	.20	.31	.41	.57

Experiment 5: The USGA is interested in determining if the four leading brands of root growth stimulator generate different mean root growth. A test area on a golf course is selected and twenty-four homogeneous plots of grass of equal size were identified. The four brands were then applied to six randomly selected plots of grass each. After two months of treatment, three cores of a uniform size were randomly selected from each of the 24 plots. The cores were then analyzed in a random order for root weight in grams. The cores were selected to assess the amount of variability in root growth over the plots. A large number of cores were misplaced and the experiment became unbalanced in that there are not six plots per stimulator and there are not three cores per plot.

Stimulator	Plots 1	2	3	4	5	6	Mean $\bar{y}_{i..}$
S1	3.2,3.3,3.4	3.1,3.5,3.0	3.2,3.1,3.4	3.4,3.0,3.1	3.3,3.3,3.1	.	3.227
	3.3	3.2	3.233	3.167	3.233	.	
S2	3.7,3.6,3.9	3.5,3.8,3.9	3.6,3.4,3.8	3.5,3.8	3.6,3.7,3.6	3.5,3.9	3.675
	3.733	3.733	3.6	3.65	3.633	3.7	
S3	3.7,3.8,3.9	3.6,3.7,3.8	3.4,3.5	3.6,3.7	3.5,3.9	3.4,3.7	3.657
	3.8	3.7	3.45	3.65	3.7	3.55	
S4	4.2,4.2,4.3	4.1,3.9,3.8	4.2,4.1,3.9	3.8,4.0,4.1	.	.	4.05
	4.233	3.933	4.067	3.967	.	.	