MA4505 Lecture 5A Design of Experiments

An **experiment** is a process or study that results in the collection of data. The results of experiments are not known in advance. Usually, statistical experiments are conducted in situations in which researchers can manipulate the conditions of the experiment and can control the factors that are irrelevant to the research objectives.

For example, a rental car company compares the tread wear of four brands of tires, while also controlling for the type of car, speed, road surface, weather, and driver.

Experimental design is the process of planning a study to meet specified objectives.

Planning an experiment properly is very important in order to ensure that the right type of data and a sufficient sample size and power are available to answer the research questions of interest as clearly and efficiently as possible, with consideration for the amount of resources available to carry out such experiments.

Designing an Experiment
Perform the following steps when designing an experiment:
1. Define the problem and the questions to be addressed.
2. Define the population of interest.
3. Determine the need for sampling.
4. Define the experimental design.

The subject matter for steps 1 to 3 were covered in Science Maths 3 (MA4603). We will look at step 4 specifically in this module.

Define the Experimental Design

A clear definition of the details of the experiment makes the desired statistical analyses possible, and almost always improves the usefulness of the results.

The overall data collection and analysis plan considers how the experimental factors, both controlled and uncontrolled, fit together into a model that will meet the specific objectives of the experiment and satisfy the practical constraints of time and money.

The data collection and analysis plan provides the maximum amount of information that is relevant to a problem by using the available resources most efficiently.

Understanding how the relevant variables fit into the design structure indicates whether the appropriate data will be collected in a way that permits an objective analysis that leads to valid inferences with respect to the stated problem.

The desired result is to produce a layout of the design along with an explanation of its structure and the necessary statistical analyses. The data collection protocol documents the details of the experiment such as the data definition, the structure of the design, the method of data collection, and the type of analyses to be applied to the data.

Defining the experimental design consists of the following steps:
1. Identify the experimental unit.
2. Identify the types of variables.
3. Define the treatment structure.
4. Define the design structure.

Terminology

(Remark: Categorical Variables are commonly called "Factors" in Statistics)

A *factor* is any aspect of the experimental conditions which may affect the result obtained form an experiment.

- A Controlled Factor is any factor that can be altered by the experimenter at will.

- Uncontrolled Factor is any factor that can't be freely altered.

- Factor Levels are the discretized values of indicating the degree of presence of a given factor (for example, high and low).

Factors

The two independent variables in a two-way ANOVA are called factors.

The idea is that there are two variables, factors, which affect the dependent variable.

Each factor will have two or more *levels* within it, and the degrees of freedom for each factor is one less than the number of levels.

Types of Variables

A data collection plan considers how four important variables: background, constant, uncontrollable, and primary, fit into the study.

In experimental design, variables are known as **factors**.

Inconclusive results are likely to result if any of these classifications are not adequately defined. It is important to consider all the relevant variables (even those variables that might, at first, appear to be unnecessary) before the final data collection plan is approved in order to maximize confidence in the final results.

Primary variables are independent variables that are possible sources of variation in the response. These variables comprise the treatment and design structures and are referred to as **Primary factors**.

Uncontrollable factors are those variables that are known to exist, but conditions prevent them from being manipulated, or it is very difficult (due to cost or physical constraints) to measure them.

The experimental error is due to the influential effects of uncontrollable variables, which will result in less precise evaluations of the effects of the primary and background variables. The design of the experiment should eliminate or control these types of variables as much as possible in order to increase confidence in the final results.

Treatment Structure

The treatment structure consists of factors that the researcher wants to study and about which the researcher will make inferences.

The **primary factors** are controlled by the researcher and are expected to show the effects of greatest interest on the response variable(s).

The levels of greatest interest should be clearly defined for each **primary factor**. The levels of the primary factors should fully represent scope of your study.

There may also be a secondary factor that you want to account for.

The levels of the primary factors can represent the entire range of possibilities or a random sub-set.

It is also important to recognize and define when combinations of levels of two or more treatment factors are illogical or unlikely to exist.

The treatment structure relates to the objectives of the experiment and the type of data that's available. *One-way, two-way, three-way, 2^n, 3^n, D-optimal, central composite*, and *two-way with some controls* are examples of treatment structures that are used to define how data are collected.

Levels of a treatment (examples)

Treatments are administered to experimental units by **level**, where level implies amount or magnitude. For example, if the experimental units were given 5mg, 10mg, 15mg of a medication, those amounts would be three levels of the treatment.

Level is also used for categorical variables, such as Drugs A, B, and C, where the three are different kinds of drug, not different amounts of the same thing.

Types of Effects

An effect is a change in the response due to a change in a factor level. There are different types of effects (See Above)

One objective of an experiment is to determine if there are significant differences in the responses across levels of a treatment (a fixed effect) or any *interaction* between the treatment levels.

Main Effects

This is the simple effect of a factor on a dependent variable. It is the effect of the factor alone averaged across the levels of other factors.

Example

A cholesterol reduction clinic has two diets and one exercise regime. It was found that exercise alone was effective, and diet alone was effective in reducing cholesterol levels (main effect of exercise and main effect of diet).

Also, for those patients who didn't exercise, the two diets worked equally well (main effect of diet); those who followed diet A and exercised got the benefits of both (main effect of diet A and main effect of exercise).

However, it was found that those patients who followed diet B and exercised got the benefits of both plus a bonus, an *interaction effect* (main effect of diet B, main effect of exercise plus an interaction effect).

Main Effect
- The main effect involves the independent variables separately.
- The interaction is ignored for this part. Just the rows or just the columns are used, not mixed.
- This is the part which is similar to the one-way analysis of variance.

Interaction Effect
- The interaction effect is the effect that one factor has on the other factor.
- The "*degrees of freedom*" is the product of the two degrees of freedom for both factors in a Two-Way ANOVA procedure.
- **Important:** The interaction effect can only be computed in the case of multiple measurements per treatment group.

If this is always the case, the analysis is usually easily manageable, given that the anomalies in the data are minimal (outliers, missing data, homogeneous variances, unbalanced sample sizes, and so on).

Replication

If a treatment condition appears more than one time, it is defined to be replicated. True replication refers to responses that are treated in the same way.

Replication is essential for estimating **experimental error**. The type of replication that's possible for a data collection plan represents how the error terms should be estimated.

Two or more measurements should be taken from each experimental unit at each combination of conditions, if possible.

In addition, it is desirable to have measurements taken at a later period in order to test for repeatability over time. The first method of replication gives an estimate of pure error, that is, the ability of the experimental units to provide similar results under identical experimental conditions.

Design Structure : Blocks

Most experimental designs require experimental units to be allocated to treatments either randomly or randomly with constraints, as in blocked designs.

Blocks are groups of experimental units that are formed to be as homogeneous as possible with respect to the block characteristics.

> You may consider blocks as natural subgroupings that would arise naturally, i.e. without intervention.

The term block comes from the agricultural heritage of experimental design where a large block of land was selected for the various "treatments", that had uniform soil, drainage, sunlight, and other important physical characteristics.

Homogeneous clusters improve the comparison of treatments by randomly allocating levels of the treatments within each block.

Design Structure : Completely Randomized Design

Subjects are assigned to treatments completely at random.

For example, in an clinical trial study, volunteer patients are randomly assigned to one of four treatment groups (three new types of a treatment and the standard).

Suppose the total number of patients in the 4 groups is 96 (i.e. randomly assign 1/4 of them, or 24 patients, to each of the 4 types of treatments).

Test Method			
Standard	New Treatment 1	New Treatment 2	New Treatment 3
24 Patients	24 Patients	24 Patients	24 Patients

Randomized Complete Block Design.

Subjects are first divided into blocks according to some natural grouping (i.e. location, demographic or medical characteristics).

Then subjects in each block are then randomly assigned to treatments so that all treatment levels appear in each block.

For example, in the clinical study, the patients would presumably have different demographic characteristics. For the sake of simplicity, let's say there are four main demographic subgroups A,B,C and D.

Patients within each subgroup are randomly assigned to one of the four types of treatments.

There might be significant variability between the subjects in each demographic subgroup, each of which contains 24 patients. Randomly assign 6 patients to each of the three types of tests and the standard.

The demographic subgroup is now the 'block'. The primary interest is in the main effect of the test.

	Treatment Method			
Subgroups	Standard	Treatment 1	Treatment 2	Treatment 3
Group A	6 Patients	6 Patients	6 Patients	6 Patients
Group B	6 Patients	6 Patients	6 Patients	6 Patients
Group C	6 Patients	6 Patients	6 Patients	6 Patients
Group D	6 Patients	6 Patients	6 Patients	6 Patients

The improvement of this design over a completely randomized design enables you to make comparisons among treatments after removing the effects of a confounding variable, in this case, different subgroups.

Example

A researcher is carrying out a study of the effectiveness of four different skin creams for the treatment of a certain skin disease. He has eighty subjects and plans to divide them into 4 treatment groups of twenty subjects each.

Using a randomised blocks design, the subjects are assessed and put in blocks of four according to how severe their skin condition is; the four most severe cases are the first block, the next four most severe cases are the second block, and so on to the twentieth block. The four members of each block are then randomly assigned, one to each of the four treatment groups.

	Treatment Method			
Subgroups	Cream A	Cream B	Cream C	Cream D
Group 1	Patient 3	Patient 4	Patient 1	Patient 2
Group 2	Patient 8	Patient 7	Patient 6	Patient 5
Group 2	Patient 11	Patient 12	Patient 9	Patient 10
….	….	….	….	….

Blinding

In a medical experiment, the comparison of treatments may be distorted if the patient, the person administering the treatment and those evaluating it know which treatment is being allocated. It is therefore necessary to ensure that the patient and/or the person administering the treatment and/or the trial evaluators are 'blind to' (don't know) which treatment is allocated to whom.

Sometimes the experimental set-up of a clinical trial is referred to as double-blind, that is, neither the patient nor those treating and evaluating their condition are aware (they are 'blind' as to) which treatment a particular patient is allocated. A double-blind study is the most scientifically acceptable option.

Sometimes however, a double-blind study is impossible, for example in surgery. It might still be important though to have a single-blind trial in which the patient only is unaware of the treatment received, or in other instances, it may be important to have blinded evaluation.

Confounding

Confounding occurs when the experimental controls do not allow the experimenter to reasonably eliminate plausible alternative explanations for an observed relationship between independent and dependent variables.

Consider this example.

- A drug manufacturer tests a new cold medicine with 200 volunteer subjects - 100 men and 100 women.
- The men receive the drug, and the women do not.
- At the end of the test period, the men report fewer colds.
- This experiment implements no controls at all. As a result, many variables are confounded, and it is impossible to say whether the drug was effective.
- For example, gender is confounded with drug use. Perhaps, men are less vulnerable to the particular cold virus circulating during the experiment, and the new medicine had no effect at all. Or perhaps the men experienced a placebo effect.

This experiment could be strengthened with a few controls. Women and men could be randomly assigned to treatments. One treatment could receive a placebo, with blinding.

Then, if the treatment group (i.e., the group getting the medicine) had sufficiently fewer colds than the control group, it would be reasonable to conclude that the medicine was effective in preventing colds.