Experimental Design

Victoria Goodall
Nelson Mandela Metropolitan University

GSSA Research Skills Workshop
Wilderness
4th July 2016

Experimental Design

Workshop Agenda

- Introduction
- Randomized experiments
- Measurement
- Design of the experiment
- Methods for data collection
- Terminology
- Components of experiment design
- Experimental designs
- Sample size
- Summary and check list

Introduction

Nelson Mandela Metropolitan University
for tomorrow

- What is Experimental Design?
  - The design of a suitable experiment to test a hypothesis often requires some ingenuity and a suspicious mind.
  - The experiment may involve very sophisticated equipment and high costs.
  - There are a number of features common to all good experiments which exist whatever the technical details.

Introduction

- The goal of any experiment is to obtain a clear answer to the research question(s), using a minimum amount of resources.
- The goal of experimental design is to attain maximum information, precision and accuracy in the results through efficient utilisation of resources.
- The design of the experiment needs to allow you to address your research question(s) of interest.
- Clearly defined objectives are the most important to an appropriate design.

Introduction

- Discrimination
  - Experiments should be capable of discriminating clearly between different hypotheses.
  - It often turns out that two or more hypotheses give indistinguishable results when tested by poorly-designed experiments.

Introduction

- Replication and Generality

    - Biological data are notoriously variable.

    - Experiments must be repeated enough times for the results to be analysed statistically.

    - Due to biological variability, generalising the results can have major drawbacks.

    - The purpose of replication is to estimate variation in response to a treatment due to unknown differences in the system

Introduction

Nelson Mandela Metropolitan University
for tomorrow

- Controls

  - The experiment must be well controlled.

  - We must eliminate by proper checks the possibility that other factors in the overall test situation produce the effect we are observing, rather than the factor we are interested in.

  - Method of measurement must be reproducible from day to day.

Randomized Experiments

- A randomized experiment is typically conducted for one or more of the following reasons:

  - to compare the response achieved at different settings of the controllable variables.
  - to determine the principle cause of variation in a measured response.
  - to find the conditions that give rise to maximum or minimum response.
  - to obtain a mathematical model in order to predict future response.

Randomized Experiments

- The distinguishing features of a randomized comparative experiment include the following:
  - conditions are controlled and manipulated by the investigator.
  - treatments are assigned at ***random*** to the experimental units.
  - inferential cause and effect relationships ***can*** be inferred.

Measurement

- When you make measurements, it is important you know both the *accuracy* and the *precision* of your measuring system.
- These two terms are not synonymous.
- Accuracy means the ability of the method to give an unbiased answer on average.
- Precision is an index of the method's reproducibility.
- Ideally your method should be both accurate (i.e., give the true mean) and precise (i.e., have a low standard deviation). Sometimes one is more important than the other.

Measurement

- Accuracy and precision together help you to judge the reliability of your data.
- The number of significant figures you should quote your results to is also important.
- For example, if you use a balance reading to the nearest gram, you should give the results to the nearest gram and not, say, to the nearest tenth of a gram.

Design of the Experiment

- Request from the statistician…
  - Think about the statistical tests you will use to analyse your data BEFORE you conduct the experiment.
  - Knowledge of the statistical test will help to design the experiment properly.

- Three steps in Experimental Design
  1. Define the objectives
  2. Devise a strategy
  3. Operational Details

Methods for Data Collection

- Non-experimental research
  - Data dredging
  - Meta-analysis

- Experimental Evidence
  - Controlled Observations
  - Manipulative Experiments

Terminology

- Experimental Unit
  - the physical entity or subject that is exposed to the experimental treatment ***independently*** of other units.

- Replication
  - an independent repetition of the experiment, in which an experimental unit upon exposure to the treatment constitutes a single replication of the experiment.

- Treatment
  - the set of circumstances created for the experiment in response to the research hypotheses.

Terminology

- Factor
  - a particular group of treatments.

- Factor levels
  - the categories of the treatment factor.

- Fixed Factor
  - the levels of the factor are assumed to be fixed (reproducible).
  - inference is restricted to those levels of the factor that were included in the studied.
  - the parameters of interest are the ***treatment effects***.

Terminology

- Random Factor
  - the levels of the factor are assumed to represent a random sample of the potential levels within a well defined population.
  - inference is directed to the population.
  - the parameters of interest are the ***<u>variance components</u>***.

Components of Experiment Design

- Treatment Structure
  - The set of treatments or treatment combinations that the investigator has selected to study and/or compare.
- Design Structure
  - The grouping of experimental units into homogenous groups; often referred to as the blocking criterion.
  - Also known as ‘Stratification’.
- Randomization
  - The process of randomly assigning the treatments to the experimental units of the design structure.

Treatment Structure

- The treatment structure should be dictated by the research hypotheses, and constructed based on those factors to be compared as measured by their effect on the given response variable.

- The treatment structure can be a set of individual treatments, referred to as a one-way treatment structure, or a set of treatment combinations, referred to as a factorial treatment structure.

- The treatment structure can also be a hierarchical arrangement involving multiple size experiment units, in which the treatment levels of one or more factors occur within the levels of one or more of the remaining factors.

Design Structure

- The design structure of an experiment involves the grouping of experimental units such that the variability of the units within the groups is less than that among all units prior to grouping. The goal is to group the experimental units in such a way that the conditions under which the treatments are observed are as uniform as possible.

- The design structure can range from a completely randomized design structure with no blocking criterion, to design structure which include multiple blocking criteria.

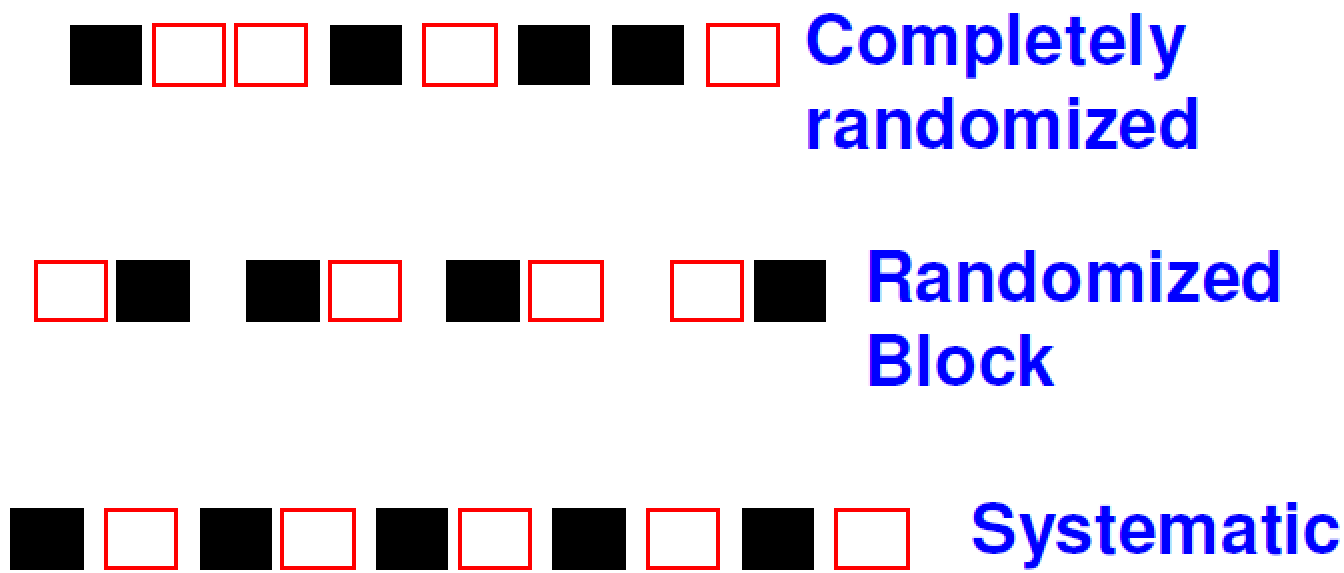
Randomization

- Randomization is the link between the treatment structure and the design structure.
- Randomization is the process of randomly assigning the treatments of the treatment structure to the experimental units of the design structure.
- Randomization functions to prevent systematic and personal biases from being introduced into the experiment.
- Randomization also functions to provide a valid estimate of error variance for justifiable statistical inference.

Experimental Designs

Nelson Mandela Metropolitan University
for tomorrow

Some acceptable Experimental Designs



Experimental Designs

Nelson Mandela Metropolitan University for tomorrow

- Single Factor Designs
- Factorial Designs
- Randomized Block and Incomplete Block Designs
- Latin Square Designs
- Split Plot Designs

Single Factor Design

- A study design with only one independent factor.
- The factor is manipulated at multiple levels.
- Often used to determine the effect of a certain treatment or intervention.
- Common analysis: One-way ANOVA

Treatment 1	Treatment 2	…	Treatment K
X_{11}	X_{12}	$\cdots$	X_{1k}
X_{21}	X_{22}	$\cdots$	X_{2k}
$\cdots$	$\cdots$	$\cdots$	$\cdots$
X_{n1}	X_{n2}	$\cdots$	X_{nk}

Factorial Design

- The comparison of the effects of two or more factors acting simultaneously on a common response or criterion variable.

- Allows for the examination of the effects of the interaction of each factor combination on the variable of interest.

Factorial Design

Factor B		Factor A			
		Level 1	Level 2	…	Level p
	Level 1	X_{111} X_{112} $\cdots$ X_{11n}	X_{121} X_{122} $\cdots$ X_{12n}		X_{1p1} X_{1p2} $\cdots$ X_{1pn}
	Level 2	X_{211} X_{212} $\cdots$ X_{21n}	X_{221} X_{222} $\cdots$ X_{22n}		X_{2p1} X_{2p2} $\cdots$ X_{2pn}
	Level r	X_{r11} X_{r12} $\cdots$ X_{r1n}	X_{r21} X_{r22} $\cdots$ X_{r2n}		X_{rp1} X_{rp2} $\cdots$ X_{rpn}

Randomized Block Design

- The experimenter divides subjects into subgroups called *blocks*, such that the variability ***within*** blocks is less than the variability ***between*** blocks.

- The subjects within each block are then randomly assigned to the treatment conditions.

Gender		**Treatment**	
		Placebo	Vaccine
	Male	250	250
	Female	250	250

Latin Square Design

- The Latin Square design is used where the researcher desires to control the variation in an experiment that is related to rows and columns in the field.

- Treatments are assigned at random within rows and columns, with each treatment once per row and once per column.

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

Split-Plot Design

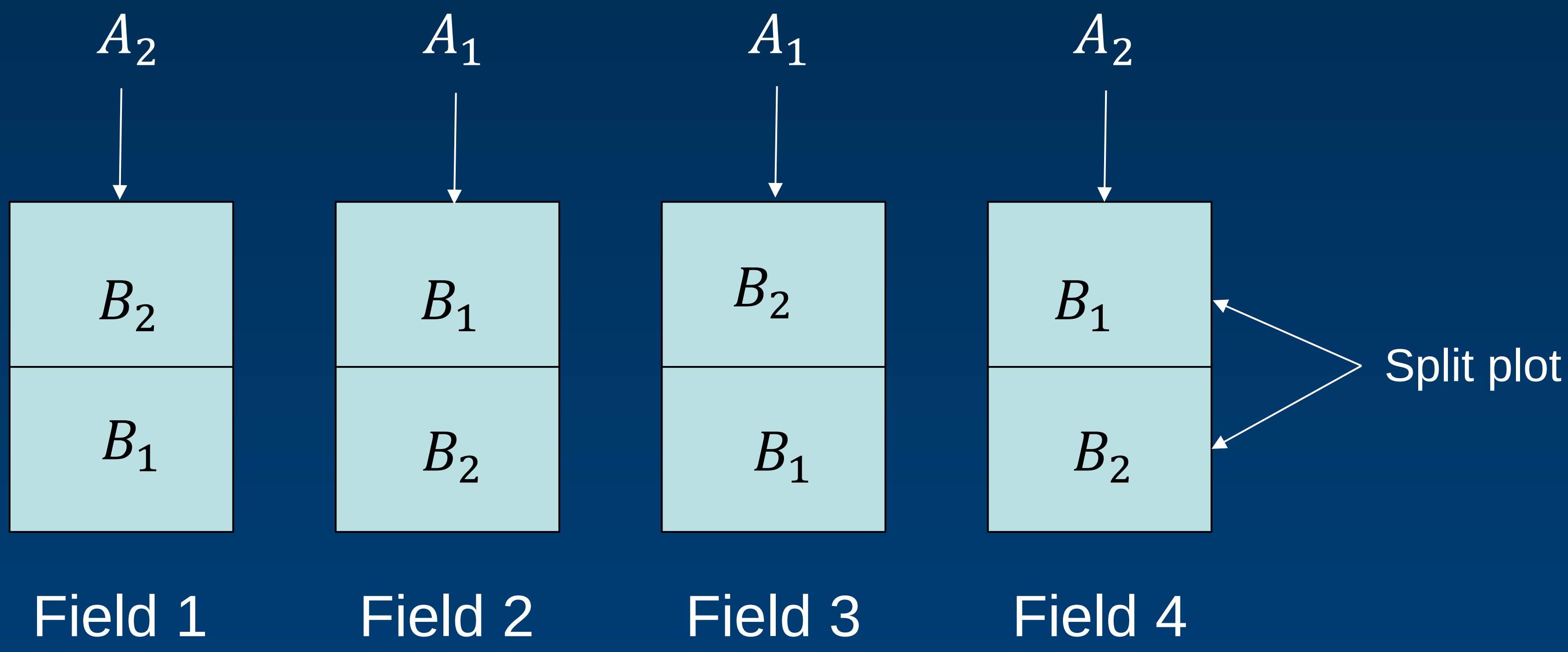
- A split-plot experiment is a blocked experiment, where the blocks themselves serve as experimental units for a subset of the factors.

- There are two levels of experimental units.

- The blocks are referred to as whole plots, while the experimental units within blocks are called split plots, split units or subplots.

- Two randomizations: one to determine the assignment of block-level treatments to whole plots, and one of treatments to split-plot experimental units within each block or whole plot.

Split-Plot Design - Example



- Factor A is whole-plot factor and Factor B is split-plot factor.

Sample Size

- The primary information required to estimate sample size:
  - an estimate of variance σ^2.
  - the size of the difference between treatment means that has physical significance.
  - the significance level of the test (α), or the probability of a Type I error.
  - the power of the test $(1-\beta)$, or the probability of a Type II error.

Hypothesis Testing Errors

- Type I error – An incorrect rejection of a true null hypothesis.

- Type II error – failure to reject a false null hypothesis.

Sample Size

- The primary information required to estimate sample size:
  - an estimate of variance σ^2.
  - the size of the difference between treatment means that has physical significant (critical difference).
  - the significance level of the test (α), or the probability of a Type I error.
  - the power of the test $(1-\beta)$, or the probability of a Type II error.

Sample Size

$$r = 2\left[z_{\alpha/2} + z_{\beta}\right]^2 \left(\frac{\sigma}{\delta}\right)^2$$

- r is the number of replications per treatment.
- $z_{\alpha/2}$ is the standard normal variate exceeded with probability $\alpha/2$.
- z_{β} is the standard normal variate exceeded with probability β.
- σ is the standard deviation of the response.
- δ is the critical difference between the two treatment means.

Sample Size

- Threats to Statistical validity
  - Low statistical power due to low sample size.
  - Violation of the assumptions of the statistical tests.
  - Use of unreliable measures.

Data

Source of Confusion	Solution
Change over time	Control Treatments
Procedure effects	Control Treatments
Experimenter Bias	Randomise treatments and controls
Initial or inherent variability among experimental units	Replication; interspersion of treatments

Uncertainty

- An estimate is of no value without some statement of the uncertainty in the estimate.

- Replications are used to quantify this uncertainty / variability.

Other comments

- Measurements made by people can be influenced by unconscious biases.
- Ideally measurements should be made *without* the knowledge of the treatment applied.
- Are the subjects really representative of the population you want to study?

Summary

Nelson Mandela Metropolitan University for tomorrow

- Good experimental design is the key to good science.
- Researchers should use experience, information from the literature and common sense to decide which factors are worth studying.
- To decide on an appropriate scale, you should consider what area, population, time period, physical, chemical or social situation you want the results to reflect.
- Consult with experts at the start of the project to discuss the design and the desired analysis.

Experimental Design Check List

a) Define the objectives of the experiment.

b) Identify all sources of variation, including:
   - treatment factors and their levels.
   - experimental units.
   - blocking factors, noise factors, and covariates.

c) Choose an appropriate rule for assigning the experimental units to the treatments.

d) Specify the measurements to be made.

e) Run a pilot study if possible.

f) Specify the model.

g) Outline the analysis.

h) Calculate the number of observations that need to be taken.


Thank You!
Questions???