The Theory of Designed Experiments

1. Basic Principles

UNESP-Botucatu, August 2010

What is this course about?

1. The Theory of Designed *Experiments*. An experiment (as opposed to an observational study or a survey) involves changing the system under study and observing the effect changes have on the system.

Why experiment?

- Observed relationships allow causal inferences to be drawn, rather than just noting patterns.
- Informative events can be made to happen.
- We can get the data we need.

What is this course about?

2. The Theory of *Designed* Experiments. Planning experiments involves applying the statistical principles of design of experiments to the specific practical situation. Essentially it is about how to conduct the experiment to collect informative data. Design and analysis are inextricably linked. All statistical modelling should pay attention to the way in which the study was designed. Likewise, the design should take account of the proposed analysis.

3. The *Theory* of Designed Experiments. Statistical theory is the set of logical principles which allow us to develop methodology for particular applications.

What is this course about?

We plan experiments to get *valid* and *efficient* answers to the questions the experiment is intended to answer.

- Validity implies quantities are estimated free from bias - accuracy.
- Efficiency implies quantities are estimated with low variance - precision.

What is this course about?

Planning experiments is not about *lists* of "experimental designs". Many textbooks and manuals for packages describe design of experiments as if there is a finite list of designs, one of which must be used.

- Mead (1988) described this as the "Procrustean approach" to design.
- Bailey (1998, *JRSS D*, 261-271) described it as "butterfly collecting".

Instead, we should "design for the experiment, not experiment for the design".

The positive approach

1. A set of experimental treatments is chosen to meet the objectives of the experiment. *Treatments* are the things which are fixed by the experimenter to be compared.
2. The experimental units are identified. *Units* are the things to which treatments are applied.
3. It is decided which response variable(s) to measure on each unit (and perhaps how often or where).
4. Any likely patterns of variation among the units are noted.
5. Any restrictions on which treatments can be applied to which units are noted.
6. Given 1-5 a method is chosen to decide which treatment is to be applied to which unit.

Model

The basic model which will (almost) always be assumed is

$$y_{i(r)} = u_i + t_r,$$

$y_{i(r)}$ is the response on unit i if treatment r is applied,
u_i is the response on unit i when a treatment with zero effect is applied,
t_r is the effect of treatment r.

This is often written

$$y_{i(r)} = \mu + t_r + e_i,$$

μ is the mean of u_i across the units,
e_i is the deviation of unit i from this mean, $\sum e_i = 0$.

Note:

- The *only* assumption is additivity of treatment and unit effects.
- $y_{i(r)}$ may represent a transformation of the measured response to ensure additivity.
- Further assumptions *might* be made about e_i, e.g. $e_i = \epsilon_i$, with $\epsilon_i \sim N(0, \sigma^2)$.
- t_r might be expressed in a form which reflects the treatment structure and the contrasts of interest, e.g. $t_r = f_s + g_t + (fg)_{st}$. Further assumptions *might* be made, e.g. $(fg)_{st} = 0$.

Experimental Units

Example

A research project was conducted to discover what affects the quality of potted plants after sale to the customer. Two "shelf-life rooms" are available, which can be temperature controlled. In a particular experiment, what is an experimental unit?

A unit is the object to which a treatment is applied, i.e. the whole unit is subject to the same treatment and different units might receive different treatments.

The definition of units depends on the treatments, and hence on the objectives, on practicalities, on material available, etc.

Example

Consider a number of different objectives:

- to discover effect of temperature;
- to discover effect of light level;
- to discover differences between varieties;
- to discover the effect of light levels at different temperatures;
- to discover differences between varieties at different temperatures and light levels.

Example

Often there is some free choice - how many plants in a plot?

- More allows better estimation of within plot variation, but might mean fewer units overall.
- Fewer might allow more experimental units overall, but allows less good estimation of within plot variation.

Note - the number of experimental units is not in general the same as the number of observations.

To compare treatments we need to separate their effects from all other sources of variation which could affect the comparison between the units to which they are applied.

Experimental Units

Other examples

- A clinical trial to compare two anti-hypertensive drugs. The blood pressure of each patient can be measured every month for four months. Is a patient a unit or 4 units?
- A field experiment to compare six varieties of wheat. A 200m $\times$ 150m field is available. How is it divided into plots?
- A laboratory experiment to compare doses of insecticide. How many insects constitute a unit?
- An industrial experiment to optimize production of a liquid product. How much production constitutes a run?

Experimental Units

It is not always obvious what should be used as units. More units might increase *representation bias*.

Example: An experiment to compare 4 hormone treatments in cattle. 16 cows available, in 4 pens.

- Applying hormones to animals individually gives replication, applying hormones to groups of animals in the same pen does not.
- In farming practice, animals in the same pen would always receive the same hormone treatment.

Response Variables

A decision will be made about what to measure from each unit, again depending on the objectives. The responses might be:

- multivariate - number of leaves, leaf colour, etc.
- repeated measurements in time - measure leaf colour every week for 6 weeks.
- spatially structured measurements - measure leaf colour from top, middle and lower leaves.

Sampling Issues

Questions of sampling design, as well as experimental design, arise in experiments, e.g.

- how to sample leaves to test colour (sampling within units);
- how to select plants for the experiment (sampling of units from a sampling frame of possible units).

Sampling Issues

Often the most important issue in sampling is to balance representativeness against practicality, e.g.

- Clinical trials - we want patients to represent target population, but we can only use a few centres, patients who are not too severely ill, etc.
- Toxicology - we want genetically diverse animals, but we want a genetic structure which will show effects.
- On-farm experiments - we want a range of environmental conditions, but cannot use too many sites or seasons.
- Industrial experiments - we want to represent normal production conditions, but it might be difficult to experiment at night.

Protecting Units

Standard analyses assume results from different units are independent. We might have to protect units to ensure this, e.g.

- In human or animal experiments, washout periods might be allowed between treatments.
- In field experiments, guard rows might be used, or only the centre of plots included in the analysis.
- In laboratory experiments, equipment must be thoroughly cleaned between runs.
- In industrial experiments, feedback control systems must be switched off during the experiment.

If protection is impossible or expensive we might model the interference instead of trying to prevent it.