ST 512 - Design of Experiments Supplement

Readings: Chapter 2 (for 2.2/2.3 read if interested)

This class is about analyzing data. As scientists, most of the time this data will come from a designed experiment, but the methods used for analysis are also useful for observational studies. However, the conclusions drawn will differ! Let's define what me mean by experimental and observational study.

Observational Study researchers does not interfere or intervene in the process of collecting data.

- Ex: measuring political beliefs in using a poll, measuring yield of a crop based on rainfall

Experimental Study researchers manipulate the conditions in which the experiment is done.

- Ex: assigning different fertilizers and irrigation method and measuring crop yield, assigning temperatures of water to tanks containing a fish and observing weight gain

Big difference in conclusions drawn!

- Cannot usually infer causation from observational experiments, but you can from a well-designed experiment.

- Experiments are not always feasible or ethical. i.e. cannot assign people to smoke a pack a day or have expectant mothers drink a certain amount of alcohol.

To describe the methods for creating a well-designed experiment, we first need some definitions.

- **Response Variable** - Variable of interest that characterizes performance or behavior.

- **Explanatory Variables** - Variables that determine the study conditions (can be quantitative or categorical).

- **Factor** - Explanatory variable of interest.

- **Level** - The specified value of a factor (or explanatory variable).

- **Confounding Variable** - Explanatory variable (not of interest) that may mask (or enhance) the effect of a factor.

- **Covariate** - Quantitative confounding variable.

- **Treatment** - A specific experimental condition, either the level of a factor (if only 1 factor) or the combinations of levels from a number of factors.

- **Experimental units (EUs)** - Units on which the treatments are assigned.

- **Measurement (Observational) units** - Units on which values are observed (often the same as EUs, but not always).

- **Replicate** - Name given to EUs that receive the same treatment.

- **Control Treatment** - Benchmark treatment sometimes necessary for comparison (to avoid the *placebo effect*).

- **Experimental Error** - Used to describe the variation in response among EUs that are assigned the same treatment.

Example: An experiment was run to determine the effects of adding phosphorous $(0, 147, 294, 441\ kg/m^2)$ and nitrogen $(0, 45, 90, 135\ kg/m^2)$ to the soil of a certain type of grass (a Miscanthus species). The growth of the plant was of interest and at the end of the growing period the plant was dried and the weight recorded with the final measurement being recorded in megagram per hectare $(0.1\ kg/m^2)$. Four plots of grass were used in total. Within each plot, each combination of phosphorous and nitrogen was observed. The plots were arranged in a large field in a 4x1 rectangle (north to south). There is a possibility of a water gradient as a stream runs to the north of the field. A partial data table is given here:

Plot	P	N	Dry yield
1	0	135	1.95
1	0	45	3.51
1	0	90	2.87
1	0	0	2.88
1	294	45	2.37
1	294	0	3.5
1	294	135	3.55
1	294	90	4.4
...	...	...	...

Let's identify (if possible) the response, explanatory variable(s), factor(s), level(s), confounding variable(s), treatment(s), number of replicates, experimental units, and observational units.
Response: Y = Growth of plant
Explanatory Variable(s): Phosphorous, Nitrogen, Plot
Factor(s): Phosphorous, Nitrogen
Level(s): For Phosphorous - (0, 147, 294, 441), For Nitrogen - (0, 45, 90, 135)
Confounding Variable(s): Plot
Treatment(s): Combos of Phos and Nitrogen Levels - (0,0), (0,45), ..., (441,90),(441,135)
# of replicates: 4
EUs: Areas of the plots that were assigned the treatments
OUs: Same as EUs here.
 (See example 2.4/2.7 and the resulting discussion for more practice)

Notice that many of the response values are different. What is causing them to be different?

Sources of Variation in the responses of an experiment:

1. **Treatment effect** - we hope there is an effect due to the variables we are setting

2. **Identified confounding variables** - We record some variables that are not of interest, but we think may have an effect on the response.

3. **Unidentified sources (these make up the Experimental Error or error variation)** -

   (a) Inherent variability in experimental units - Experimental units are different!
   Ex: No two people, paper towels, concrete blocks, or even lab rats are exactly the same.
   Consequence: Experimental units respond differently to the same treatment

   (b) Measurement error - Multiple measurements of a same experimental unit typically contain error.
   If the same experimental unit is measured more than once, will the value be the same?
   Ex: Blood Pressure, Break a water sample in two, measure each for bacteria

   (c) Variations in applying/creating treatments
   The treatment is not clearly defined, leaving room for interpretation.
   Ex: Two researchers mix concrete, one stirs for 10 minutes and one for 20 minutes, will they come out exactly the same? Temperature is of interest but two ovens don't heat exactly the same, etc.

   (d) Effects from any other extraneous (or lurking) variables - Extraneous variables are those variables that are not part of the treatment, but may influence the response.
   Ex: For the oven example, the experiment is done over the course of several days. There may be slight differences due to humidity changes.

Let's identify these in the grass growth example.
(a) Each area of grass might differ causing the responses to differ
(b) When drying the grass for weighing, maybe you lose some or the scale varies
(c) For some plots the Phos is added first then the Nitrogen, for others they Phos and Nitrogen are mixed then added. This may cause a difference in response
(d) Could be many things such as sun variance across plots or neighboring area differs (trees for other grass), etc.
 (See example 2.5/2.6 and the resulting discussion for more practice)

No matter how hard we try, some experimental error will remain. What we can do is use good experimental design techniques to ensure our study is valid.

DOE is about creating the optimal experiment to determine the effects of different treatments. Different types of experimental designs are then analyzed differently.

Pillars of Experimental Design

1. Randomization - means that the treatments are randomly allocated to the EUs.

   (a) Every EU has a chance to get a different treatment, so helps protect the results of the analysis against a systematic influence of lurking variables.

   (b) Allows the observed responses to be regarded as a random sample.

   Note: Different randomization schemes lead to different statistical analyses.

   Completely Randomized Design - for t treatments, replicated n_t times each, use a random number generator to assign the treatments to the EUs.

   Most basic randomization design - assumes all EUs are exchangeable.

2. Replication - Repetition of an experiment using a large group of subjects to reduce chance variation in the results

   (a) Allows us to generalize the results to the population and increases reliability of conclusions.

   (b) Allows an estimate of variability (an estimate of experimental error) not due to the treatment effect.

   Note: Replication does not mean that we measure the same EUs multiple times, this is called repeated measures. Observations from repeated measures experiments cannot usually be considered independent.

3. **Methods for accounting for/reducing experimental error**

   (a) Controlling Variables - holding certain variables constant across the EUs
       Decreases generalizability, but reduces experimental error.

       We're not interested in the effects of these variables on the response. These variables affect the response in exactly the same manner, so that we don't see the effects on the conclusions. We don't get information on what happens at levels other than the fixed one.

   (b) Blocking - Divide subjects with similar characteristics into 'blocks', and then within each block, randomly assign subjects to treatment groups.

       Blocks - Groups of EUs sharing a common level of a confounding variable.
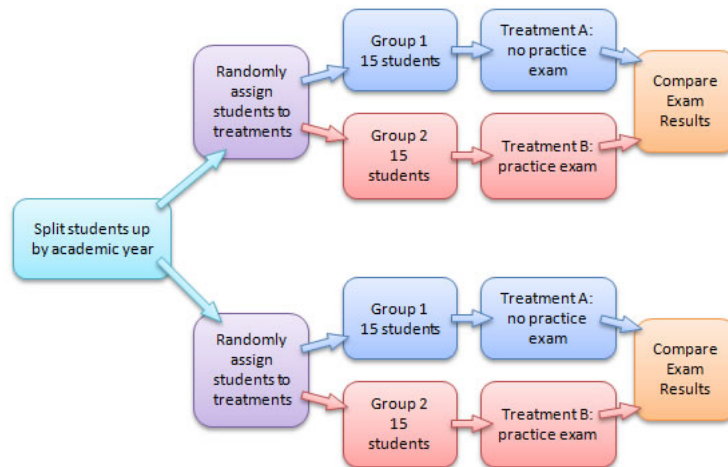
       Similar to controlling, but allows for increased generalizability. EUs within a block are very similar (decreases experimental error there as all the EUs in a block are affected similarly by the confounding variable). By having enough blocks to cover the range of the population you can still generalize.)

There are also methods for dealing with some explained experimental error during the analysis stage - Namely ANCOVA.

These ideas are very important. Unless you are well versed in statistical methods and ideas you should consult a statistician before investing time and money in an experiment.
'A poorly designed study can never be saved, but a poorly analyzed one has the possibility of being reanalyzed.'