Introduction to Design of Experiments

Professor: Hammou El Barmi
Columbia University

- Researchers use experiments to answer questions. Typical questions might be:
  1. Is a drug a safe and an effective cure for a disease?
  2. This could be a test of how AZT affects the progress of AIDS.
  3. How will long-distance telephone usage change if our company offers a different rate structure to our customers?
  4. Does short-term incarceration of spouse abusers deter future assaults?
- Example: Consider the spousal assault example mentioned above. Justice officials need to know how they can reduce or delay the recurrence of spousal assault. They are investigating three different actions in response to spousal assaults. assailant could be warned, sent to counseling but not booked on charges, or arrested for assault. Which of these actions works best? How can they compare the effects of the three actions?

- The following few chapters deal with comparative experiments.
- We wish to compare some treatments.
- We compare treatments by using them and comparing the outcomes. Specifically, we apply the treatments to experimental units (EU) and then measure one or more responses.
- For the spousal assault example, the treatments are the three actions by the police. Individuals who assault their spouses could be the experimental units, and the response could be the length of time until recurrence of assault.
- We compare treatments by comparing the responses obtained from the experimental units in the different treatment groups.
- This could tell us if there are any differences in responses between the treatments, what the estimated sizes of those differences are, which treatment has the greatest estimated delay until recurrence, and so on.

- A key feature in designed experiments is that the experimenter holds the power of deciding which treatment gets assigned to which experimental unit.
- Having that control allows us to make stronger inferences about the nature of differences that we see in the experiment. Specifically, we may make inferences about causation.
- This last point distinguishes an experiment from an observational study. An observational study also has treatments, units, and responses. However, in the observational study we merely observe which units are in which treatment groups; we don't get to control that assignment.
- The drawback of observational studies is that the grouping into "treatments" is not under the control of the experimenter and its mechanism is usually unknown. Thus observed differences in responses between treatment groups could very well be due to these other hidden mechanisms, rather than the treatments themselves.

Component of an experiment

An experiment has treatments, experimental units, responses, and a method to assign treatments to units.

- Treatments are the different procedures we want to compare. These could be different kinds or amounts of fertilizer in agronomy, different long-distance rate structures in marketing, or different temperatures in a reactor vessel in chemical engineering.
- Experimental units are the things to which we apply the treatments. These could be plots of land receiving fertilizer, groups of customers receiving different rate structures, or batches of feedstock processed at different temperatures.

- Responses are outcomes that we observe after applying a treatment to an experimental unit. That is, the response is what we measure to judge what happened in the experiment; we often have more than one response. Responses for the above examples might be nitrogen content or biomass of corn plants, profit by customer group, or yield and quality of the product per ton of raw material.
- Randomization is the use of a known, understood probabilistic mechanism for the assignment of treatments to units. Other aspects of an experiment can also be randomized: for example, the order in which units are evaluated for their responses.
- Types of randomization (not exhaustive):
  1. Unrestricted: For each EU, assign a treatment according to a discrete uniform distribution of the treatments. Can result in unequally replicated treatments.
  2. Restricted to ensure equal replication (e.g. 10 are to receive each drug)
  3. Restricted to ensure equal replication within selected subgroups of EUs. e.g. Say we have 20 people of which 10 are male and 10 are female. We can restrict randomization so that 5 males and 5 females are in each of the two treatment groups.

How to actually carry out the randomization-some ways

- Label the EUs from 1 to N where N is a multiple of the number of treatments, g. Put these on slips of paper and draw them from a hat, on at a time, w/o replacement. Assign first to N/g labels the first treatment, next N/g to second treatment, etc.
- Generate a list of labeled N random numbers from a U(0,1) distribution and rank them in increasing order. Assign the first N/g labels to treatment 1, etc. E.g. N=8, g=4

0.187	0.502	0.380	0.968	0.887	0.894	0.066	0.681
1	2	2	4	3	4	1	3

Component of an experiment

- Why randomize?
- Randomization provides protection against confounding
- Confounding occurs when the effect of one factor or treatment cannot be distinguished from that of another factor or treatment.
- Example: Consider planting corn variety A in Minnesota and corn variety B in Iowa. In this experiment, we cannot distinguish location effects from variety effects?the variety factor and the location factor are confounded

Component of an experiment

- Experimental Error is the random variation present in all experimental results.
- Different experimental units will give different responses to the same treatment, and it is often true that applying the same treatment over and over again to the same unit will result in different responses in different trials.
- Experimental error does not refer to conducting the wrong experiment or dropping test tubes.

Control has different uses in design.

- First, an experiment is controlled because we as experimenters assign treatments to experimental units. Otherwise, we would have an observational study.
- Second, a control treatment is a standard treatment that is used as a baseline or basis of comparison for the other treatments. This control treatment might be the treatment in common use, or it might be a null treatment (no treatment at all).