STAT 5200 Handout #3: Statistical Terminology (Ch. 1-2)

Common Terminology in Design of Experiments

Treatments
These are the procedures that we want to evaluate. For example:

- Different drugs or therapies in medical trials.
- Different strains of feed, or amounts of fertilizer or irrigation in agricultural experiments.
- Baking temperatures for a cake product.

Factors
Usually we group related treatments into a single factor. For example, in an agricultural experiment

- One factor may comprise 3 different amounts ("treatments") of fertilizer.
- Another factor may be 4 different amounts of irrigation.
- A third factor could be 2 harvesting regimes.

Levels
The treatments that make up a factor are called the levels of the factor.

- Sometimes there will be an ordinal structure to the levels of a factor. E.g., amounts of fertilizer or irrigation.
- Other times the levels of a factor may be quite distinct drugs, or chemicals, or therapies, etc. For example, comparing several different kinds of pain reliever.
- Some experiments may have both kinds of factors.

Experimental Units
The individual "entities" (subjects, plots, objects, groups, etc.) that are assigned to and that receive individual application of a factor level (or combination of factor levels).

- A human subject is assigned a certain pill.
- A pen full of young turkeys is assigned a certain diet.
- An acre plot in a field is assigned a certain irrigation amount from a single sprinkler line.
- A class full of students is assigned a certain pedagogical intervention.
- A single batch of cookies is baked at a certain temperature and with a certain type of butter.

Sometimes called true replicates

drives statistical power (how likely it will be to detect true effects)
(higher → better)

Sample Size
This usually refers to the total number of experimental units.

Response Variable or Outcome (often more than one in an experiment)
This is what we measure on each experimental unit. We are interested in identifying how the factor(s) affect(s) these measurements. For example,

- Blood pressure in humans
- Mortality rate for young turkeys
- Yield of a crop
- Average score on a final exam
- Average hardness of cookies

Measurement Unit (sometimes called pseudo-replicates)
This is the unit that measurements of the response variable are actually made on.

- Often the measurement units will be exactly the same as the experimental unit.
- Sometimes a measurement unit is a subset of an experimental unit. For example:
    - Apply the diet to whole pens of about 250 turkeys, but only measure response on a random sample of 30 turkeys from each pen.
    - Expose whole class of students to pedagogical intervention, and record each individual student's final exam score.
    - Bake whole batch of cookies, and measure hardness of each individual cookie.

Usually take average of msmnt units within each exp. unit (otherwise, risk false claims overstating sample size)

Control
This has two very different meanings in the context of experimental design.

- [Verb] One use of the word refers to the fact that the experimenter gets to allocate the experimental units.
- [Noun] In many experiments, including clinical trials with people, new treatments are compared with a standard treatment or no treatment at all. Control is the term that is used to refer to that standard treatment or no treatment.

Blind and Double Blind
These terms are most commonly used in the context of medical experiments

- An experiment involving humans in which the subjects are not told which treatment they received is called ***blind***.
- If the persons evaluating the experimental subjects at the end of the experiment also do not know which treatment each subject received then the experiment is said to be ***double blind***.

These are very desirable characteristics for experiments involving humans. They can

- Eliminate conscious or unconscious bias from the subjects self-reporting.
- Ensure that adherence rates are essentially the same for all treatment groups, including the control.
- Eliminate conscious or unconscious bias of the persons evaluating the experimental subjects.

Placebo → To maintain the blind, so only difference between trt. & ctl. groups is treatment itself

- In order to make an experiment involving people in which one of the treatments is no treatment (a control) blind, subjects that receive the control are given a fake treatment called a ***placebo.***
- The placebo resembles the treatments being administered to the other subjects as much as possible. For example,
  - A saline injection.
  - A pill made or sugar and flour
- Sometimes placebos may be very elaborate, as in gastric freezing:
  - Ulcer patients swallowed a balloon with a freezing solution, and generally reported substantial pain relief
  - In a follow-up, some patients received freezing solution while others received body temp. solution, with respective improvement rates of 34% and 38%.

In experiments involving pain medications:

- About 1/3 of persons who receive a placebo report pain relief of some kind.
- Bigger placebo pills work better than little placebo pills at reducing pain.
- Brightly colored placebos work better than drab colored placebos at reducing pain.

In experiments that do not involve people it is still often desirable that the persons evaluating the samples do not know which treatment each experimental unit received.

- Such experiments are called ***blind***.
- Blinding eliminates investigator bias, conscious or unconscious.
- For example, the homeopathy ("water memory") experiment published in *Nature* (1988)

→ Evaluators should be blinded if any subjectivity is possible

Experimental Error (also known as ***chance error***) → "difference among repeats" (not mistakes)

- If an experiment is replicated in every way except with new experimental units, then one would expect slightly different results each time.
- One of the objectives of experimental design is to reduce the size of experimental error as much as possible for a fixed number of experimental units.

Confounding (cannot be undone!)

This is what happens when the effects of two or more factors cannot be distinguished. For example,

- Different corn varieties grown in different states. V1 - IA V2 - IN V3 - IL
- Different kinds of feed always placed at the same place in a field. F1 - edge F2 - center

Careful experimental design can eliminate confounding.

Allocation of Experimental Units

How should one allocate experimental units to the treatments (factor levels)?

- Using some probabilistic mechanism. E.g., coin tossing, rolling a die, or (usually) using computer generated random numbers.
- Packages such as SAS have programs for doing this for you. (Example coming up in Handout #4) (make reproducible by setting the seed)
- We say that an experiment in which the experimental units have been assigned by some probabilistic mechanism is <u>randomized</u>. (NOTE: haphazard ≠ random)

Why Randomize?

- Randomization protects against experimenter bias in the allocation of experimental units.
- Randomization protects against confounding of unmeasured factors and factors of interest.
- (*) Randomization is the basis for ALL statistical inference carried out on experimental data. E.g., hypothesis tests for treatment effects.
- An experiment without randomization is a poor experiment and cannot be analyzed statistically.

– because of possible confounding factors

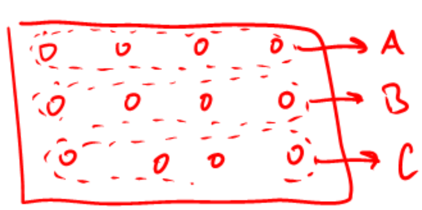
process of drawing population conclusion from sample data (hypothesis test)

How to Randomize?

- Randomization methods vary for different experimental designs.
- In many experiments there are constraints on the randomization, including requirements for each treatment to occur the same number of times in each day or field or block. Sometimes randomization is carried out in 2 or more stages (e.g., split plot designs). Examples of such experimental designs are coming up in Unit E (Ch. 13 & 16).

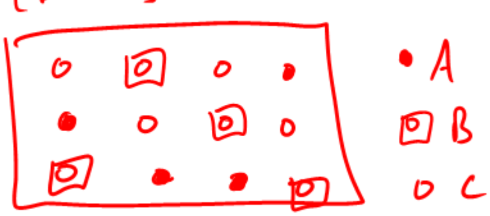
Example: field with 12 trees (3 rows, 4 columns); effects of 3 sprays are to be compared

- Design 1: [Actual]



· Experimental unit: row (N=3)
· Measurement unit: tree
· application is to each row

- Design 2: [Better]



· Application to indiv. trees
→ exp. unit = msmnt. unit = tree (N = 12)