Stats! Almost Didn't Make it….

Unit 2 - Sample Experiments

Your Late Professor Colton

Unit 2 - Outline

<u>Experiments, Good and Bad</u>

- Explanatory vs Response
- Subjects, Treatments and Control
- Lurking and Confounding Variables
- Placebo and Blinded Experiments
- Randomized Comparative Designs
- Principles in Experimental Design
- Statistical Significance

<u>Experiments in the Real World</u>

- Simple Designs
- Completely Randomized Design
- Matched Design and Matched Pairs Design
- Block Design
- Generalizing and Other Issues
- Application

Review: Experiments vs. Observational Studies

Experiments

- ***Imposes*** treatments and controls ***randomly*** to groups.
- Can help determine a causation.
- Reduces the potential a *lurking variable* can affect the results (controls for them).
- More accurate to determine a relationship between the explanatory variable and response.

Observational Studies

- **Observe** what happens without imposing restraints.
- No random assignment.
- Can reveal an association or correlation between variables but not causation.

Experiments, Good and Bad: Explanatory vs Response Variables

Explanatory Variable

- The variable(s) being used to **explain** a response (the variable doing the explaining).

Response Variable

- The variable we are ultimately interested in (the one trying to be explained).

Our GOAL is to see if we can describe our response with an explanatory variable(s).

Experiments, Good and Bad: Subjects, Treatment and Control

<u>Subjects</u>

- **Subjects** (also called **Experimental Units**) are what the *treatment* is being *applied* to.

<u>Treatment</u>

- **Treatment** is the *experimental condition.*

<u>Different Groups in an Experiment</u>

- **Treatment Groups**
    - The group the receives the treatment.
- **Control Group**
    - The group that does not receive the treatment.
    - We have a control group to compare our treatment groups to.
    - We have to know what the “base” level of our response is.

Experiments, Good and Bad: Lurking vs Confounding Variables

Lurking Variable

- A **Lurking Variable** is a variable that is *not* among the explanatory or response variables in a study but *may influence the response variable*.
  - *Lurking = **not** included*

- Exist in all studies, we can't measure everything.
  - Can cause problems in establishing cause-and-effect relationship between response and explanatory.

Confounding Variable

- A **Confounding Variable** is a variable whose effects on the response variable *cannot be distinguished* from one or more of the *explanatory variables* in the study.
  - *Confounding = included*

- With well designed studies, we can take into account some confounding variables
  - A blocking variable is a confounding variable that is being controlled.

Experiments, Good and Bad: Lurking vs Confounding Variables

<u>Lurking Variable EXAMPLE</u>

- **Setup**: We believe that ice cream sales can explain the increase in shark attacks.
- **Question**: Is there really a direct cause and effect here???
- Of course not! In the summer, more people are outside at the beach enjoying their banana splitz.
- More people in the ocean for the sharks… Time of the year (or weather) is a lurking variable. Obviously this affects the response and it's not something we measured in our simple silly experiment.

<u>Confounding Variable EXAMPLE</u>

- **Setup**: We plant tomatoes in a garden that is half-shaded.
  - We test a fertilizer by putting it on the plants in the sun and apply none to the shaded plants.
  - Months later the fertilized plants bear better and more tomatoes.
- **Question**: Is it the fertilizer, the sun, or the combination of the two that leads to a better crop???
- The effect of the fertilizer and the effect of the sunshine are confounded.

Experiments, Good and Bad: Placebo and Blinded Experiments

Placebo

- A fake treatment.
- In pharmaceutical studies, a placebo would have no active ingredients.

Placebo Effect

- Sometimes subjects respond favorably to a treatment due to the expectation of a cure.
  - This effect is called the **placebo effect**.
- The placebo effect can be *confounded* with the *effect of a treatment*.
  - In that case, the researcher *cannot distinguish which effect*, treatment, or placebo effect influenced the patient responses.

Blinded Experiments

- **Single Blind**: *Subjects* don’t know which group they are in.
- **Double Blind**: *Subjects and data collectors* don’t know which group the subjects are in.
  - Helps avoid bias imposed by a doctor for example.

Experiments, Good and Bad: Placebo and Blinded Experiments

Placebo Effect Example

- Another study told 13 people who were very sensitive to poison ivy that the stuff being rubbed on one arm was poison ivy. It was a placebo, but all 13 broke out in a rash. The stuff rubbed on the other arm really was poison ivy, but the subjects were told it was harmless—and only 2 of the 13 developed a rash.
- In my sprained ankle example, even people that received a sugar pill might feel a reduction in pain because they believe they are actually getting something to help the pain.

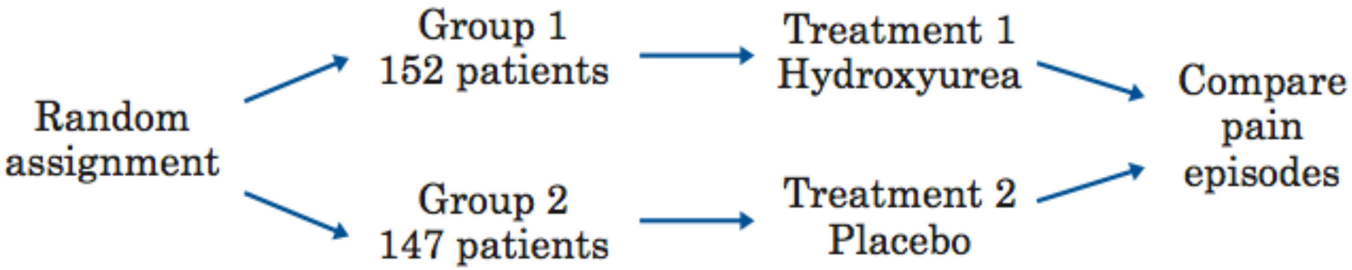
Blinded Experiments Example

- Doctors' expectations change how they interact with patients and even the way they diagnose a patient's condition. It is also foolish to tell doctors and other medical personnel what treatment each subject is receiving. If they know that a subject is getting "just a placebo," they may expect less than if they know the subject is receiving a promising experimental drug.
- Back to the example, single blind would be if I knew whether the you were taking an aspirin or a sugar pill, but you did not. Double blind would be if we both did not know.

Experiments, Good and Bad: Randomized Comparative Experiment

GOAL

- To ensure our *design* it will detect whether the explanatory variables **cause** changes in the response variables.
- Basic idea is to **compare** two or more treatments and assign the treatments **randomly** to the subjects.
- This way, other factors like *lurking and confounding variables* **impact all subjects equally**, and…
  - *Any systematic differences* in the responses of subjects receiving different treatments can be **attributed to the treatments** rather than **random chance**.
- Simple example with only two treatments (one is a control):



Experiments, Good and Bad: Logic of Experimental Design

- **Randomization** produces groups of subjects that should be similar, on average, in all respects before the treatment is applied.
- **Comparative design** exposes all groups to similar conditions, other than the treatments they receive.
  - This ensures that any additional lurking variables operate equally on all groups and, on average, groups differ only in the treatments they receive.
- Therefore, differences in the response variable are likely due to the effects of the treatments rather than random chance.

Summary: Good Design of Experiments - 4 Principles

1. Control
    a. Make conditions as similar as possible for all groups so that the only difference is the imposed treatment
2. Randomization
    a. Equalize effects of known and unknown variation.
    b. Ex. Don’t separate by gender, time of day, location.
3. Replication (use enough subjects)
    a. Apply each treatment to a number of subjects
    b. Repeat for different levels of the treatment.
4. Blocking (optional)
    a. Group by individuals that are similar together in some manner, then randomize treatments within the block.

Experiments, Good and Bad: Statistical Significance

We compare…

- Do we expect the treatments to give the exact same results?
- No, this difference could be due to the subjects chosen or the way they were assigned to treatments (chance).

Our conclusion…

- How different do the results have to be to decide if one treatment is better than another?
- An observed effect of a size that would rarely occur by chance is called **statistically significant.**

Experiments in the RW: Simple Designs

<u>Review</u>

- Sampling in practice has difficulties that just using a random sample won't solve.
- So we stratify, or cluster for example.

<u>New</u>

- Using a randomized comparative experiment is also a good idea, but it doesn't solve all the difficulties of experimenting.
- The logic of a randomized comparative experiment *assumes that all the subjects are treated alike except for the treatments* that the experiment is designed to compare…
- This is not always the case, what are some other methods then???

Experiments in the RW: Completely Randomized Design

<u>Completely Randomized Design (CRD)</u>

- Simplest design.
- In a **completely randomized experimental design**, all the experimental subjects are allocated at random among all the treatments.
    - *Just assign treatments from the entire single pool of subjects. No grouping or dividing up or pairing first.*
- A completely randomized design can have any number of explanatory variables.

Experiments in the RW: Matched Pairs Design

Matched Pairs Design

- A little more complex design, combines matching with randomization.
- A **matched pairs design** compares just two treatments.

Strategy

- Choose pairs of subjects that are as closely matched as possible.
- And then randomly assign treatment A to one subject and the other to treatment B.
- Or giving both treatments to the same subject.

Advantage

- Helps reduce effect of lurking variables.
- Conditions are most similar for your pair, so better chance it’s the treatment showing up and not something else.

EXAMPLE

- **Setup**: *Consumers Reports* describes a method for comparing the effectiveness of two insect repellents. The active ingredient in one is 15% Deet. The active ingredient in the other is oil of lemon eucalyptus.
- **Design**: For each volunteer, the left arm is sprayed with one of the repellents and the right arm with the other.
- This is a matched pairs design in which each subject compares two insect repellents.

Experiments in the RW: Block Design

Block Design

- A little more complex design again, analogous to stratified random sample.
- A **block** is a group of experimental subjects that are known before the experiment to be similar in some way that is expected to affect the response to the treatments.

Strategy

- Choose blocking variable, divide subjects into groups based on this.
- Then randomly assign treatments within each block.

Advantage

- Takes into account confounding variables.
- Helps to single out the effect of the treatment on the response.

EXAMPLE

- **Setup**: Studying effect of running on heart rate.
- **Design**: Divide participants based on age and smoking history. Then assign subjects to walk a mile, speed walk a mile, jog a mile, etc.
- Block design because we grouped (controlled) based on some variables BEFORE randomly assigning treatments.

Experiments in the RW: Overall Example

Setup: Women and men respond differently to advertising. An experiment to compare the effectiveness of three television commercials for the same product will want to look separately at the reactions of men and women, as well as assess the overall response to the ads.

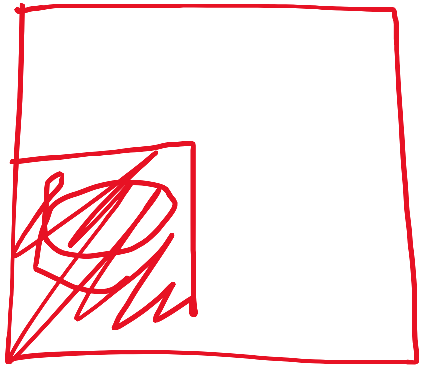
<u>CRD</u>

- A completely randomized design considers all subjects, both men and women, as a single pool.
- The randomization assigns subjects to three treatment groups without regard to their sex.
- This ignores the differences between men and women.

<u>Block Design</u>

- A better design considers women and men separately.
- Randomly assign the women to three groups, one to view each commercial. Same for the men.
- This allows us to draw separate conclusions about each block.

- Blocking also allows more precise overall conclusions, because the systematic differences between men and women can be removed when we study the overall effects of the three commercials



Summary: Matching and Blocking

Both are ways to balance lurking variables that may have an effect. Ex. location of plants.

- Matching
  - Attempts to balance treatment groups by pairing individuals by similar characteristics so that the treatment itself is the major difference between the two.
- Matched Pairs Design
  - Subjects paired with themselves
- Blocking
  - Used in experiments
  - Group similar experimental units together.
  - Treatments randomly assigned within the blocking variable to reduce variability.

Experiments in the RW: Generalizing

Generalizing

- A well-designed experiment tells us that changes in the explanatory variable cause changes in the response variable.
- Can we generalize our conclusions from our little group of subjects to a wider population?
  - The first step is to be sure that our findings are statistically significant, that they are too strong to often occur just by chance.
  - The serious threat is that the treatments, the subjects, or the environment of our experiment may not be realistic.

Example

Setup: A psychologist wants to study the effects of failure and frustration on the relationships among members of a work team.

- She forms a team of students, brings them to the psychology laboratory, and has them play a game that requires teamwork.
- The game is rigged so that they lose regularly. The psychologist observes the students through a one-way window and notes the changes in their behavior during an evening of game playing.

Questions:

- Is this scenario the “same” as spending months on an important work project to have it cancelled?
- Does the behavior of the students in the lab tell us much about the behavior of the team whose product failed?

Experiments in the RW: Other Issues

- Sample surveys suffer from nonresponse…
  - Failure to contact some people selected for the sample
  - Or refusal of others to participate.
- Experiments with human subjects suffer from similar problems.
  - Dropping out of an experiment.
  - Not following the rules.
  - These are a few things that can impact the interpretation of results the experiment.

Real World Example

Setup: The (other universities stats class) STA 261 team is interested if the amount of caffeine affects students memory of a 5 minute video shown in class. A morning class and an afternoon class were selected. Within each the class, 33 students were randomly chosen to receive one 15 mg pill of caffeine, another 33 were randomly chosen to receive a 30 mg pill of caffeine; the rest of the class (also 33 students) were given a placebo pill that had 0 mg of caffeine. Students were then scored on the amount they were able to remember on the video.

First think about what the different variables are. What is the response? Treatment? Is there a blocking variable?

Real World Application

Response Variable: *The score on the amount what students remember from the video.*

Treatment: *Amount of caffeine*

Block: *Class time*

Design diagram*:* Experimental Design Diagrams help us make sense of a design and visualize what is going on. When drawing, be very specific about what is going on. Include sample sizes, treatment names, etc.