STAT 5200 Handout #1: Experiments (Keys and Examples) (Ch. 1-2)

Experiments are used in many fields to answer many different kinds of questions:

- Is a new drug effective? Does it work better than existing drugs and other treatments?
- Which, of a number of types of grass has the highest yield, and under what conditions (of irrigation, fertilization, harvesting, etc)?
- Which combination of carbohydrate and protein sources provides the best nutrition for young turkeys?
- Does giving released prisoners small amounts of $ reduce the recidivism rate?
- What combination of ingredients leads to the strongest concrete?
- Which combination of dry ingredients, cheeses, and bake times produces the best pizza?
- Which treatments reduce the pollution levels in mud the most and in the shortest time?

Want to find some combination of "inputs" to optimize some "output"

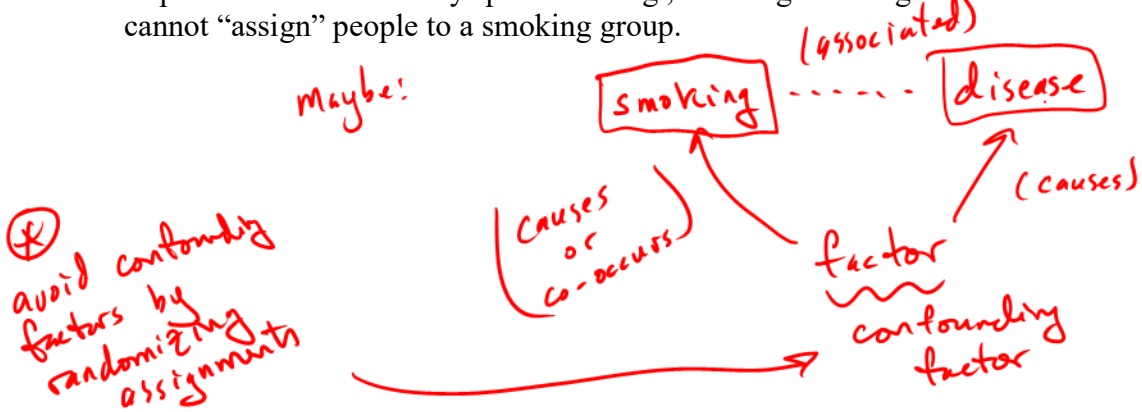
Some questions …

- What is the definition of an experiment?
  - An experiment is a study in which the experimenter controls the allocation of subjects ("experimental units") to at least one of the factors being evaluated.
- How does experimentation differ from other forms of investigation?
  - Experiments differ from other types of studies in the control that the experimenter has over the assignment of subjects to factors or treatments.
- Why is an experiment superior to other forms of investigation?
  - By appropriate allocation of subjects, one can eliminate the effects of other factors, measured and unmeasured, and can draw conclusions about causal relationships between factors and measured responses.
  -

observational studies can only show association

Pros and Cons of Experiments

- Experiments are the "Gold Standard" of the world of studies. They provide the best quality information and ***by far*** the best evidence of causality.
- Experiments are generally more expensive—often ***much*** more expensive—than other forms of investigation.
- Experiments are not always possible. E.g., smoking and lung cancer in humans. One cannot "assign" people to a smoking group.



Check: HW 0-A: #16) 1.77
#26) ± 4.57 (or F = 20.87)

Example 1: NHANES
Every 15 years or so, the National Institutes for Health conduct the National Health and Nutrition Examination Survey (NHANES). In NHANES III about 36,000 people were randomly selected, interviewed and subjected to a thorough medical examination. This survey gives us all manner of very useful information on the health and nutritional status of Americans, but is it an experiment?

No - researchers only observe people's choices/circumstances (so can't claim causation)

Example 2: The Salk Polio Vaccine, Part I
In the early 1950s the polio vaccine due to Jonas Salk looked most promising. The National Foundation for Infantile Paralysis (NFIP) proposed the following design (which was adopted by many school districts):

- Grades 1 and 3 would be the controls and receive no vaccination;
- Grade 2 would be vaccinated.

Some parents of children in grade 2 refused to allow their children to be vaccinated, so there were children in grades 1, 2, and 3 who were not vaccinated.

	Size	Rate (Polio Cases per 100,000)
Grade 2 (vaccine)	225,000	25 ← could be lower in part due to age
Grades 1 and 3 (control)	725,000	54 } only diff. is age
Grade 2 (no consent)	125,000	44 }

Questions:
1. Was the NFIP study design an experiment? Yes - assignment was controlled by researchers
2. Was it a ***good*** experiment?
Not as good as if had used randomization (so potential for confounding factors)

Example 3: The Salk Polio Vaccine, Part II
Another experimental design that was used by many school districts was the following. Parents of children in grades 1—3 were asked if they would consent to have their children participate in a randomized, controlled experiment in which the children would be randomly assigned to Vaccine and non-Vaccine groups. The children in the non-Vaccine group received a placebo and none of the children were informed as to which group they were in.

		Size	Rate (Polio Cases per 100,000)
	Treatment	200,000	28 } suggests vaccine works
difference here suggests behavior or health diffs.	Control	200,000	71 }
	No consent	350,000	46

Questions:
1. Was this an experiment? Yes - assignment was controlled
2. Was it a ***good*** experiment?
Yes - it used randomization to resolve confounding factors (incl. blinding-based behavioral changes)

Example 4
Many of the best studies on the relationship between smoking and various health problems such as lung cancer, emphysema, heart disease, and various other cancers involve looking at a cohort of people for which we have accurate, long-term health records. For example, veterans, nurses, doctors, etc. Rates of lung cancer, heart disease, etc. for smoker and non-smokers are then compared, usually within age groups and gender. These are good studies, but they are NOT experiments; why?

Researchers did not assign subjects to smoke or not

Example 5
An experiment is carried out to assess the effects of two treatment factors (pesticide type and years of thatch accumulation) on the number of fire ant mounds on golf greens. (Fire ant mounds are not desirable on golf greens.) Three separate locations (fields) are available for this experiment. Each field is divided into four parts, and each part is divided into three regions. Within each field, parts are randomly assigned pesticide types, and within part, regions are randomly assigned years of thatch accumulation. For example, one field is divided with treatment assignments as follows (Pesticide @ Thatch):

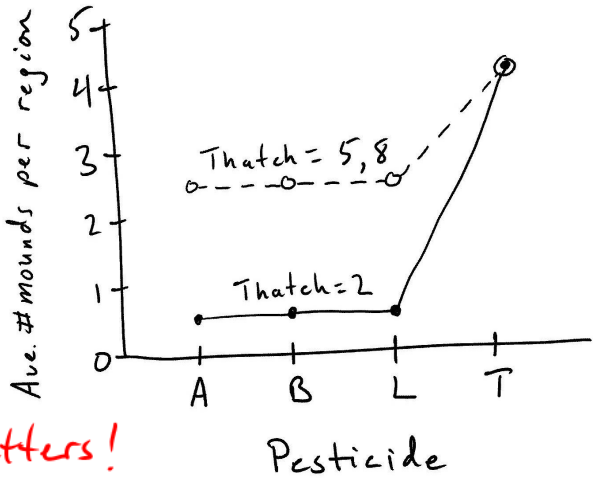
Pesticide

	Region 1	Region 2	Region 3
Part #1: Advion	Advion @ 5 yrs.	Advion @ 8 yrs.	Advion @ 2 yrs.
Part #2: Bifen	Bifen @ 8 yrs.	Bifen @ 2 yrs.	Bifen @ 5 yrs.
Part #3: Lorsban	Lorsban @ 8 yrs.	Lorsban @ 5 yrs.	Lorsban @ 2 yrs.
Part #4: Talstar	Talstar @ 2 yrs.	Talstar @ 8 yrs.	Talstar @ 5 yrs.

At the end of the study, the number of active fire ant mounds in each region is counted. From previous data, researchers suspect that the effect of years of thatch accumulation depends on pesticide type, as in the figure below.

Some issues we will consider: (later in course)

1. Why does it matter how the "treatments" were assigned?
   Randomization (constrained or nested)
2. Why do we need three fields? How much better would the study be if we could repeat this in more fields?
   Variability in repeat measurements; power / sample size calculation
3. Why does it matter that the effect of pesticide may depend on years of thatch accumulation?
   How we account for factors matters!



4. Why does the distribution of counts matter?
   Uniform? Poisson? Normal (Gaussian)?
   ↳ counts ← Negative Binomial?
   ↳ will need to assume (& check) some distribution