Experiments!

November 1, 2026

Experiments

Experiments are a scientific study where treatments are randomly *assigned* to experimental units

- Latin root: "test" or "try"
- An **Experimental Unit** is the object *a* treatment is applied to
  - Eg a car a new gasoline/ethanol blend is assigned and put into
  - Eg a farmer's field that recieves soybean hybrid A
  - Note singular: one experimental unit can have only one treatment
  - Often we can physcially touch it (safety check)
  - It's NOT the variable we are maniupulating (seed variety, gasoline blend)
- ASSIGNED is the critical word with respect to experiments
  - Without assignment it's observational

Factors, Levels, and Treatments

A **factor** is a categorical input variable (eg crop rotation type). **Levels** of a factor are the categories the variable can take.

- Eg (Corn), (Corn and Soybean), (Corn, Soy, Alfalfa), and (Corn, Soy, Winter Oats) would be the four levels of the variable "Crop Rotation"

A **treatment** (trt) is a combination of (all of the) input variables

- Eg (seed variety A) + (fertilizer type 2)
- If there is only one input the treatment and the variable are the same

	Fertilizer 1	Fertilizer 2
Corn (mono crop)	Corn + Fert 1	Corn + Fert 2
Corn-Soybean rotation	Corn/Soy + Fert 1	Corn/Soy + Fert 2
Corn-Soy-Alfalfa rotation	C/S/A + Fert 1	C/S/A + Fert 2
Corn-Soy-Winter Oats rotation	C/S/WO + Fert 1	C/S/WO + Fert 2

Experimental Units (EUs)

Experimental units are what we randomly (ideally) *assign* treatments to

- Without *assignment* it's an observational study
- Without *random assignment* it's a biased study

Observational Units is the thing we observe/measure and write down in our spreadsheets

- Common for EU's and OU's to be the same but not always
  - Eg applied seed variety A to the *entire* field at once so the entire field is the experimental unit
  - Any given soybean plant in the field is an observational unit
  - Often EU and OU differ when non-independent objects are...
    - ★ measured in time (*Repeated measures*) or
    - ★ measured in physical space (*Sub – sampling*)
- The confusion is often the impetus of *pseudo – replication*

Exp. Unit Summary

- *Experimental units* are the (usually) physical objects that treatments are assigned to
  - Experimental units should(?) be independent of each other
- *Observational units* are the physical things that we are measuring
  - Lack of (independence and identically distributed) between observation units
- The **response variable(s)** is the variable we are interested in and are measuring.
- *Treatments* are the "categories" that are being assigned during the experiment

Exp. Unit Helpful Advice

THE EASIEST WAY TO IDENTIFY THE EXPERIMENTAL UNIT IS TO THINK ABOUT WHAT THE TREATMENT IS BEING ASSIGNED TO!!

Experiments: Example

There is a field split in 64 sections outside of Ames, IA owned and operated by Iowa State University. In one study, crops of different types (corn, soybean, white oats, alfalfa) were planted and, in early spring of the following year, a sample was taken of biomass from each section of the field.

- Treatment?
- Experimental Unit?
- Observational Unit?

Experiments: Example

There is a field split in 64 sections outside of Ames, IA owned and operated by Iowa State University. In one study, crops of different types (corn, soybean, white oats, alfalfa) were planted and, in early spring of the following year, a sample was taken of biomass from each section of the field.

- Treatment? The seed varieties
- Experimental Unit? The $1/64^{th}$ section of the field
- Observational Unit? The same as the experimental unit

Experiments: Example

ISU also does research in wine making (just a more interesting crop study....). 12 rows of grape vines were in the experiment. Each row was assigned one of four fertilizers and was either aggressively trimmed or was only mildly trimmed.

At the end of the season ten bottles of wine were made from each row of grapes. Then, each bottle underwent a chemical analysis and the results (nutrition content, polyphenals, etc...) were recorded.

- Variables we are controlling (the factors)? What are their levels?
- Treatment?
- Experimental Unit?
- Observational Unit?

Experiments: Example

ISU also does research in wine making (just a more interesting crop study....). 12 rows of grape vines were in the experiment. Each row was assigned one of four fertilizers and was either aggressively trimmed or was only mildly trimmed.

At the end of the season ten bottles of wine were made from each row of grapes. Then, the wine underwent a chemical analysis and the results (nutrition content, polyphenals, etc...) were recorded.

- Variables we are controlling? What are their levels?
  - Trimming: Aggressively trimmed or mildly trimmed
  - Fertilizer: 1 of 4 types
- Treatment? A combination of trimming and fertilizer (eg Aggressively trimmed with fertilizer number 2)
- Experimental Unit? An entire row of grape vines
- Observational Unit? A single bottle of wine

Observation vs Experimental Units: Who Cares

The reason observational and experimental units need to be kept seperated is because an experimental unit is considered the foundational piece of information.

The total number of experimental units presents the upper limit of your information

At the end of the day we had 12 grape vines where each vine got a treatment. We didn't have 120 grape vines (120 = 10 bottles x 12 vines)

The ten bottles of wine give us a better understanding of that experimental unit and variability in the system

Special Types of Treatment Groups

A **control treatment group** is the treatment group that is considered "standard" and all other treatments are compared to.

- Often just called "control"
- Not always needed
- A **blind** control group is a control group who don't know they are the control group
  - **Double blind** means that neither the reseacher taking measurements nor the participants know if the participant is part of the control
- A non-blind control group would be for something like crop studies

Special Types of Control Groups

- **Placebo Group**: A **placebo** is a fake treatment used to "fool" participants
  - Often called a "sugar pill" but that's inaccurate
  - **Placebo effect** is when there is a statistically noticable change when there shouldn't be based on the fake treatment

- **Gold Standard Group**: The baseline for comparison is the best, common treatment on the open market
  - Medical studies use this most often
  - Want improvement over existing treatments

- **Non-intervention Group**: The group used for comparisons is measured but no treatment is actually *applied*
  - Pushes limit of definition of an experiment
  - Often used when we are comparing some intervention to what would happen "naturally"

Experiments: Causality?

Some people claim that experiments can prove causality (because the only expected difference in the final results is the different treatments....)

Can experiments show causation, that is that event A caused event B?

Experiments: Causality?

Some people claim that experiments can show causality (because the only expected difference in the final results is the different treatments....)

Can experiments prove causation, that is that A caused B?

No, no, no No No NO NO no NOO no a thousand times NO!

Experiments: No Causality

Claim: Experiments cannot prove causation.

Experiments: No Causality

Claim: Experiments cannot prove causation.

Arguments:

1. The analysis depends on our assumptions
   - True independence between exp units?
   - "identically distributed" There is no reason any exp unit will be different? No lurking variables leading to different means?
   - Sample Distribution $\approx$ Population Distribution?
   - Ability to generalize from our *very* restrictive experimental conditions

Experiments: No Causality

Claim: Experiments cannot prove causation.

Arguments:

1. The analysis depends on our assumptions
2. Our models are a simplification
   - Central limit theorem gets us *close* to normality
   - Measurement error, hand waving away lurking variables
   - Difference between a "straight" line and a 17^{th} order polynomial that is extremely flat?
   - Eg Did we account for relativity?

Experiments: No Causality

Claim: Experiments cannot prove causation.

Arguments:

1. The analysis depends on our assumptions
2. Our models are a simplification
3. Our estimates are wrong
   - No one truly believes there sample's estimates are population's parameter (ie the true value)
   - Sample Distribution $\approx$ Population Distribution isn't good enough, we need equality (w/ probability 1)
   - Confidence intervals wouldn't exist for experiments if we could prove causality.....

Experiments: No Causality

Claim: Experiments cannot prove causation.

Arguments:

1. The analysis depends on our assumptions
2. Our models are a simplification
3. Our estimates are wrong

In conclusion: Using estimates we expect to be wrong and assumptions we don't trust we build an overly simplified model.....and then claim to have proven causality?

Basically, Plato's Allegory of the Cave

Experiments: No Causality: Counterargument(?)

Henry Bauer in "Scientific Literacy and the Myth of the Scientific Method" argues....

- Experiments don't prove causation....but neither does anything else
- To claim causation a scientist(!!) needs to offer a valid scientific mechanism for why the causation exists
- Further, he claims part of being a scientist is making the leap from statistical uncertainty to making a solid claim
  - Eg Chargroff's rule states that, for DNA nucleotide bases, T = C and A = G
  - Statistical estimates before then had them very close to 50:50 but not exactly.

I don't consider this to be incompatible with my views outside of point 1 (which isn't experiments or my problem). Really comes down to Box's Science and Statistics working together

What do you study Professor?

What do you study Professor?

Good question! I study experimental design....

Experimental Design

We help people pick which treatments should be picked and how many replicates of each treatment should be picked.

For example imagine a situation where we apply Chemical X in differing amounts to a jungle foilage. The question is how much foilage would be destroyed given various amounts of Chemical X. We have ten simulated jungle patches we are allowed to spray.

Choices...

Two different possible experiments are below.

Trial ID	Exp 1	Exp 2
1	0	.05
2	0	.15
3	0	.25
4	0	.35
5	.5	.45
6	.5	.55
7	1	.65
8	1	.75
9	1	.85
10	1	.95

Why/when would we choose Experiment 1 vs Experiment 2?