Week 10: Experimental design

ANTH 674: Research Design & Analysis in Anthropology

Professor Andrew Du

Andrew.Du2@colostate.edu

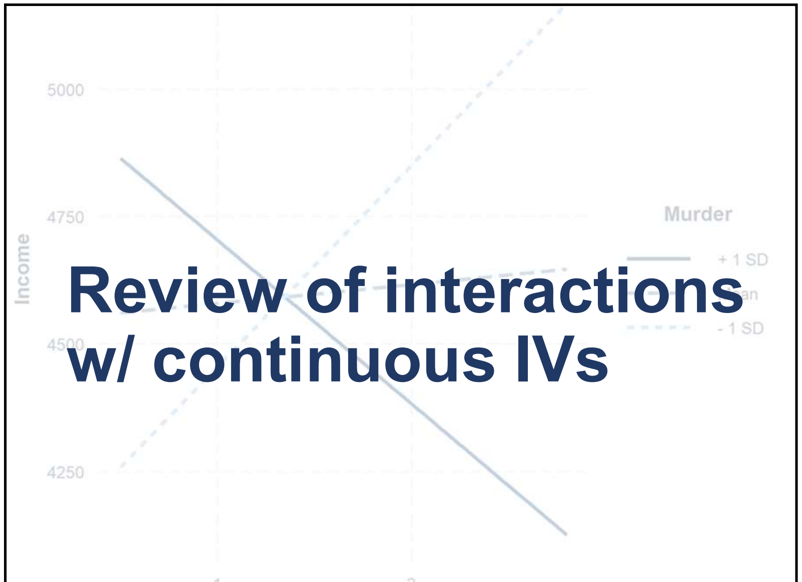
Office Hours: Thursdays, 9:00am–12:00pm
In person: GSB 312
Virtual: https://tinyurl.com/F22ANTH674

1

Lecture outline

1. Review of interactions w/ continuous IVs
2. What is experimental design?
   - Broadly used to encompass sampling design
3. How do experiments account for uncontrolled variation in DV?
   1. Replication
   2. Adequate spacing between replicates
   3. Randomization
   4. Blocking
4. Two-factor experiments
5. Moving from ANOVA to regression

2

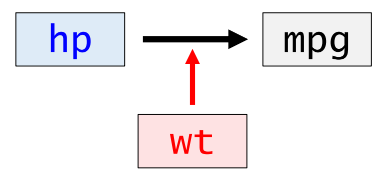


Review of interactions w/ continuous IVs

3

Interaction w/ continuous IVs

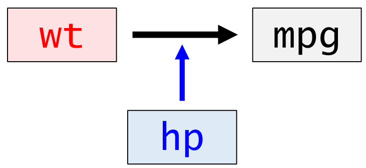
- $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \varepsilon$
- Changing intercepts **AND** slopes, just like in ANCOVA
- E.g., `mpg ~ hp * wt`
- Intercept and slope of `mpg ~ hp` will change as `wt` increases



4

Interaction w/ continuous IVs

- $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \varepsilon$
- Changing intercepts **AND** slopes, just like in ANCOVA
- E.g., `mpg ~ hp * wt`
- Intercept and slope of `mpg ~ wt` will change as `hp` increases



5

Interaction w/ continuous IVs

- `mpg = 49.81 - 0.12hp - 8.22wt + 0.03hp*wt`
- If `hp = 0`, `mpg = 49.81 – 8.22wt`
- If `hp = 1`, `mpg = 49.81 – 0.12 – 8.22wt + 0.03wt`
- If `hp = 1`, `mpg = 49.69 – 8.19wt`

6

Interaction w/ continuous IVs

- `mpg = 49.81 - 0.12hp - 8.22wt + 0.03hp*wt`
- If `wt = 0`, `mpg = 49.81 – 0.12hp`
- If `wt = 1`, `mpg = 49.81 – 8.22 – 0.12hp + 0.03hp`
- If `wt = 1`, `mpg = 41.59 – 0.09hp`

7

Interaction w/ continuous IVs

- `mpg = 49.81 - 0.12hp - 8.22wt + 0.03hp*wt`
- `49.81` is intercept when both IVs equal zero
- Focusing on `mpg ~ wt`
  - `-0.12` is change in intercept as `hp` increases by one
  - `-8.22` is slope when `hp` is zero
  - `0.03` is change in slope as `hp` increases by one

8

Interaction w/ continuous IVs

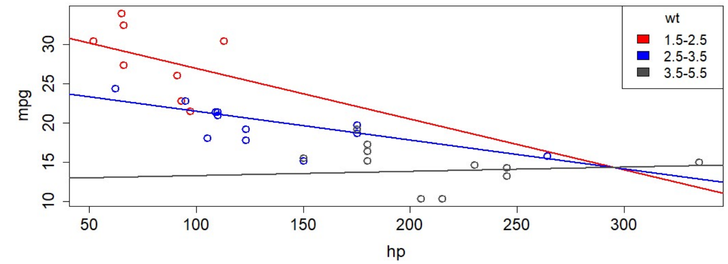
- `mpg = 49.81 - 0.12hp - 8.22wt + 0.03hp*wt`
- `49.81` is intercept when both IVs equal zero
- Focusing on `mpg ~ hp`
  - `-8.22` is change in intercept as `wt` increases by one
  - `-0.12` is slope when `wt` is zero
  - `0.03` is change in slope as `wt` increases by one

Interaction term is symmetrical!
Affects both slopes in the same way

9

Interaction w/ continuous IVs

- Focusing on `mpg ~ hp`
  - `-8.22` is change in intercept as `wt` increases by one
  - `0.03` is change in slope as `wt` increases by one



10

Questions?

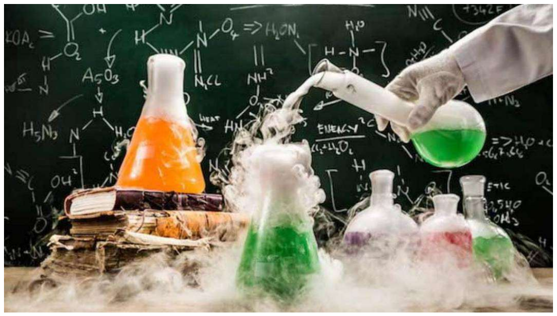
11

What is experimental design?

12

What is experimental design?

- What do you think of when I say “experiment”?



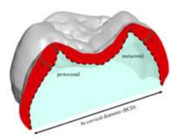
13

What is experimental design?

- Planning the logical structure of your experiment to anticipate statistical issues
- Good ED recognizes sources of unmeasured variation & plans statistics around them
  - Increases statistical power
  - Analysis of data goes hand-in-hand w/ ED
- If there are serious problems w/ ED, very difficult to correct afterwards w/ statistics
- Important to think through ED **before** data collection (clear research question helps **a lot**)!

14

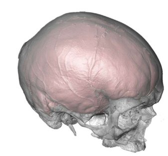
Not only experiments!



- Sampling design
- E.g., Interested in primate enamel thickness as it relates to diet
  - How many primate species to sample in museum?
  - How many individuals per species?
  - How many species or individuals per diet?
  - From what regions/years should I sample?
- Also applies to collecting data from literature if there’s a lot of data to sample from

15

Not only experiments!



- If not much data, have no choice but to use all of it, so sampling design not as applicable
- E.g., how hominin brain size evolution was affected by climate change
- Must use statistics to account for a lot of the things we’ll be talking about

Everything that follows applies to experimental **AND** sampling design!

16

Main goal of experiments

- **To see if and how one or more IVs affect a DV**
- Controls for other factors affecting DV, thereby isolating IV's effect on DV
- E.g., how does nitrogen affect plant growth rate?

Control	Treatment	
0N	+N	Good experimental design?
Sandy soil	"Regular" soil	

17

Main goal of experiments

- **To see if and how one or more IVs affect a DV**
- Controls for other factors affecting DV, thereby isolating IV's effect on DV
- E.g., How does diet affect primate enamel thickness?

Hard food	Soft food	Good "experimental" design?
Sandy soil	"Regular" soil	

Uncontrolled factor confounded w/ IV—not good!

18

Good ED accounts for:

1. Confounding factors
   - Randomization, blocking
2. Experimenter bias
   - Randomization
3. Noise and variation in data
   - Replication
4. Non-independence in data
   - Adequate spacing between replicates, randomization

Basically, **clearly** outline your research question & think about all the factors that can screw up your analyses and inference!

19

Two main types of experiments

Manipulative
- Experimenter varies IV & measures DV
- E.g., manipulating N levels → plant growth
- Expensive and time-consuming
- Good for controlling confounding factors
- Too "artificial"?

Natural
- Uses natural variation in IV & measures DV
- E.g., measure areas w/ different N levels
- Cheaper and faster
- More difficult to control confounders
- More "realistic"?

20

Questions?

21

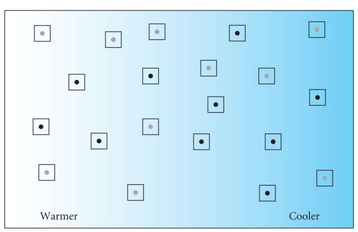
How to account for uncontrolled variation in ED

Warmer Cooler

22

ED strategies

1. Replication
2. Adequate spacing between replicates
3. Randomization
4. Blocking



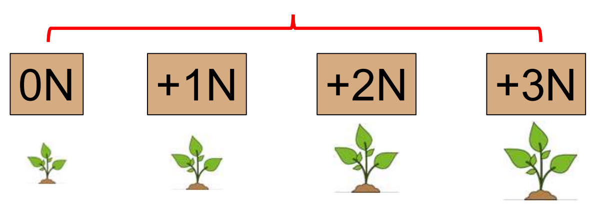
23

Basic experimental design

- One or more factors
- One or more levels w/in each factor

Interested in plant growth rate ~ N level

One factor: nitrogen; 4 levels: 0, +1, +2, +3

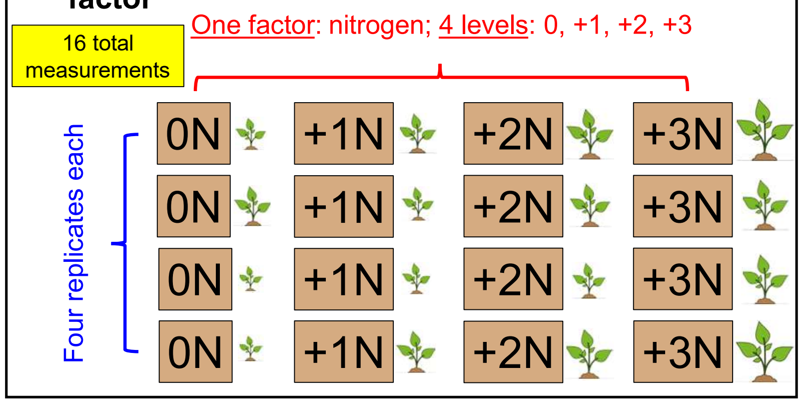


24

1. Replication

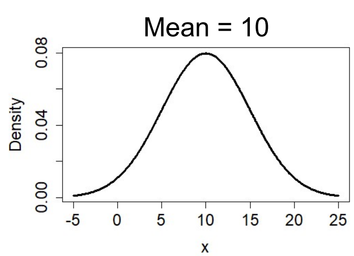
Replication

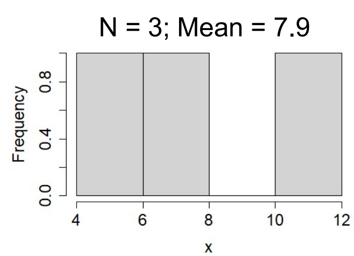
Multiple observations w/in each level of a factor



Why replicate?

- **To average out individual variation among observations w/in a treatment level**
- Same reason why a small sample may not be representative of the population



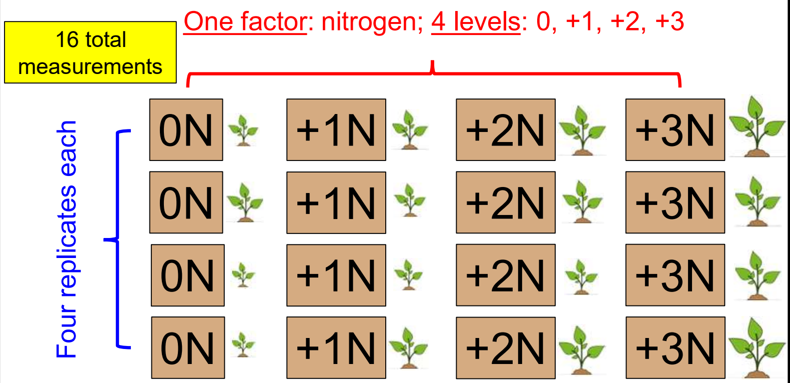


How many replicates?

- Ultimately determined by money, effort, and time, so usually there's a max # measurements
- Is a trade-off between number of factors, levels, and replicates

Replication

- If I had two levels, I could have eight replicates



29

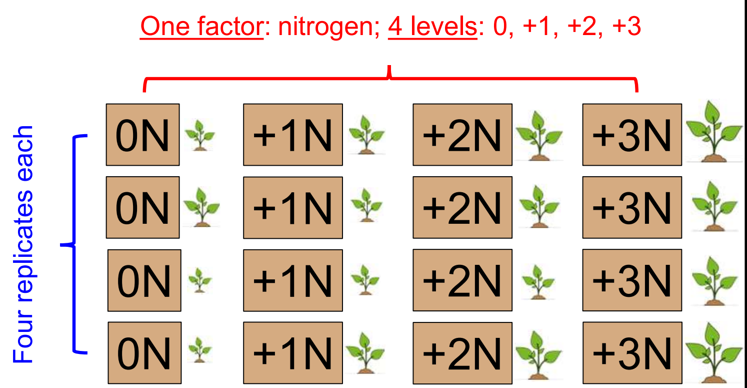
How many replicates?

- Ultimately determined by money, effort, and time, so usually there's a max # measurements
- Is a trade-off between number of factors, levels, and replicates
- Depends on your question!
  - Usually 1-3 factors is most that is manageable
  - Usually 2-5 levels is adequate
- Gotelli & Ellison's "Rule of 10": absent other information, need 10 replicates
- If data are variable and effect sizes (i.e., mean differences) are small, need more replicates

30

Which statistical test?

- This is a one-way ANOVA: `aov(plant ~ N)`



31

Questions?

32

2. Adequate spacing

33

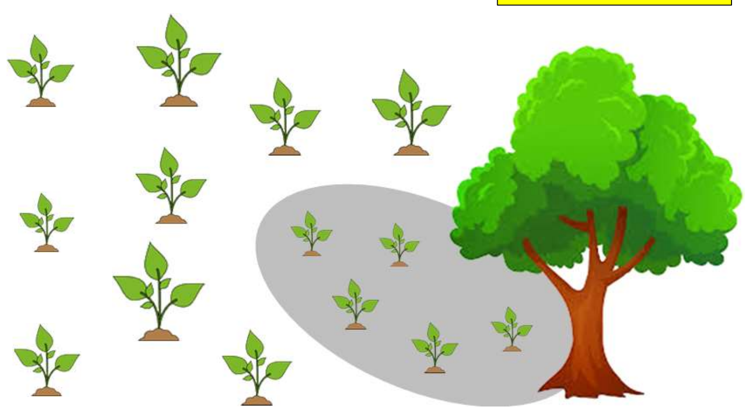
Adequate spacing

- Virtually everything is temporally, spatially, and phylogenetically autocorrelated
- Tobler's first law of geography: "everything is related to everything else, but near things are more related than distant things"
- This non-independence artificially inflates sample size, making P-values too small (increases Type I error)

34

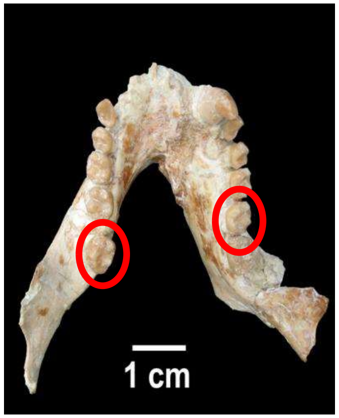
A spatial example

Seedlings in shade are smaller (size is spatially autocorrelated)



35

An extreme example



- Independent replicates?
- **NO!** Products of the same genetics, habitat, diet, & whatever processes affected this individual
- Replicates are not truly independent (i.e, **"pseudoreplication"**)

36

Adequate spacing

- Adequate spacing in time and/or space ensures replicates are truly independent
- How much time or space? Depends on process of interest
  - E.g., predation of mantis requires less spacing than lion predation
- Only worry about this if you care about P-values

37

Questions?

38

3. Randomization

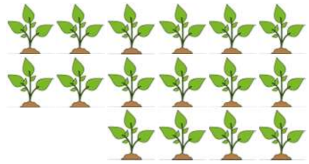
39

Randomization

- **Randomly assign replicates to treatment levels**
- Must be truly random!
  - Use random number generator, e.g., `sample()`
  - Coin flips, roll a die
- Counteracts experimenter bias, confounding factors, and non-independence
- Replication and randomization only effective when used together

40

Randomization



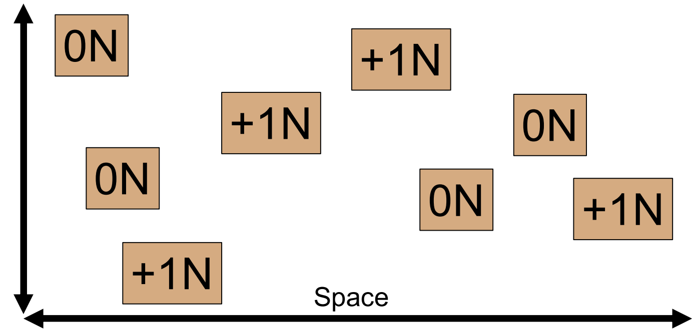
- Take individual plants and randomly assign them to treatment levels

0N	+1N	+2N	+3N
0N	+1N	+2N	+3N
0N	+1N	+2N	+3N
0N	+1N	+2N	+3N

41

Battling non-independence

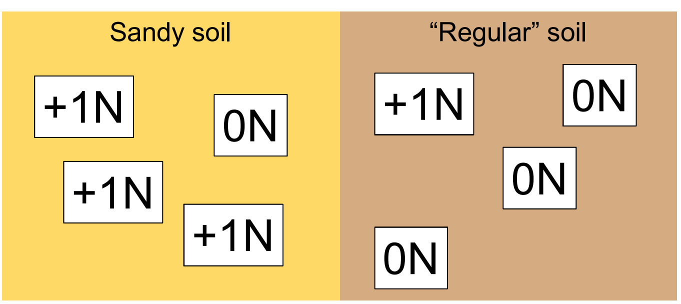
- Random distribution ensures levels that are close are canceled by levels that are far



42

Battling confounding factors

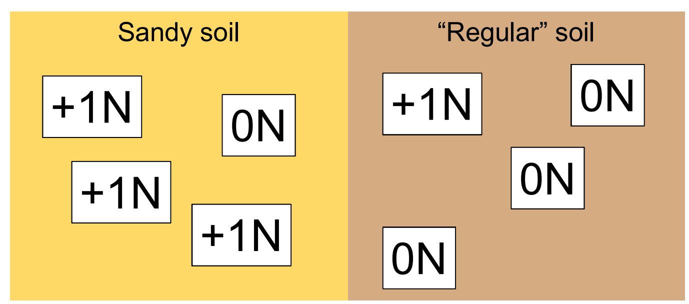
- Plots are randomly distributed
- Assigning treatments to plots is random
- Assigning individuals to treatments is random



43

Battling confounding factors

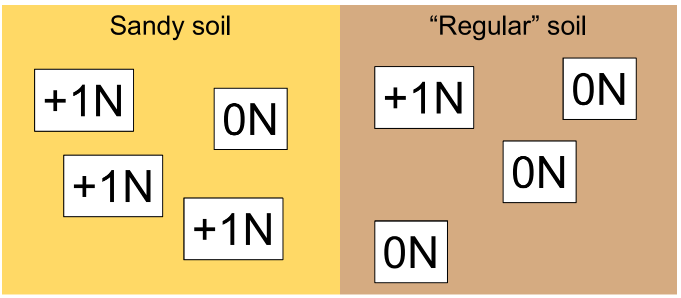
- Treatments found in both soil types now
- Treatment is now independent of confounder
- Works for "known unknown" confounders too!



44

Which statistical test?

- One-way ANOVA: `aov(plant ~ N)`
- Can include soil type as covariate to increase power (move DV variation from error term to covar.)
- Two-way ANOVA: `aov(plant ~ soil + N)`



45

Questions?

46

4. Blocking

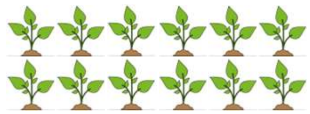
47

Blocking

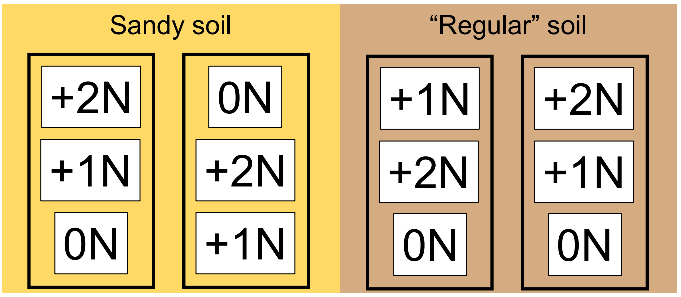
- In soil example, three "+1N" and only one "0N" in sandy soil (uneven sample sizes)
- Blocking solves this issue (**randomized block design**)
- **Block: delineated area or time period w/in which environment is homogeneous**
- Environmental variation between blocks must be greater than w/in blocks

48

Blocking



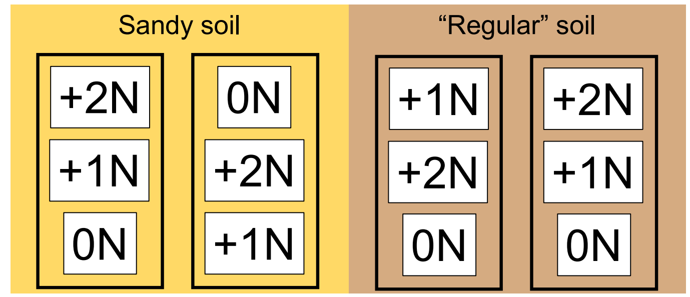
- Each treatment level assigned to a block but randomized within
- Replicates randomly assigned to levels



49

Blocking

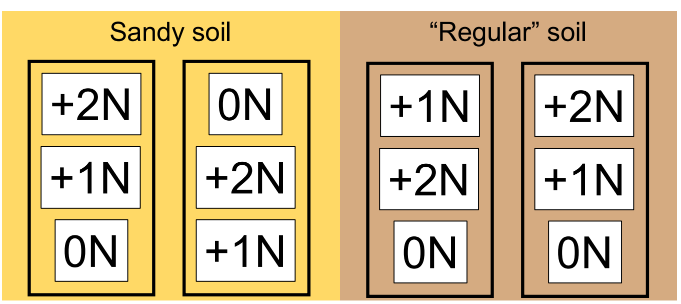
- Treatment now independent of known and unknown (spatially autocorrelated) confounders



50

Which statistical test?

- Can include block as covariate to increase power
- Two-way ANOVA: `aov(plant ~ block + N)`



51

Caveats for blocking

- Assumes that there is no interaction between blocks & treatments
  (e.g., effects of +2N > +1N > 0N for all blocks)
- What if this is not the case? Need two-factor experiment

52

Questions?

53

Two-factor experiments

54

Two-factor experiments

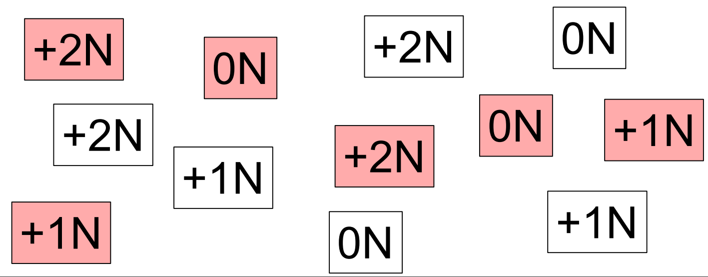
- Interested in how DV responds to manipulating **<u>TWO</u>** factors simultaneously
- Every level of factor 1 *must* be combined w/ every level of factor 2 (**<u>fully crossed</u>** or **<u>orthogonal</u>** design)
- Can look at interaction between factors 1 & 2
- With blocking, this becomes a **<u>split-plot design</u>**

55

Two-factor experiments

- E.g., Plant growth rate ~ nitrogen level **<u>AND</u>** presence of herbivores
- Plots randomly distributed, treatments randomly assigned to plots, replicates randomly assigned to plots

Herbivore present



56

Which statistical test?

- This is a two-way ANOVA w/ an interaction between the IVs (often interested in how effect of IV1 on DV is affected by IV2)
- If you have unequal # replicates among levels for the two factors, ANOVA can be problematic (http://onlinestatbook.com/2/analysis_of_variance/unequal.html) (https://mcfromnz.wordpress.com/2011/03/02/anova-type-iiiiii-ss-explained/)

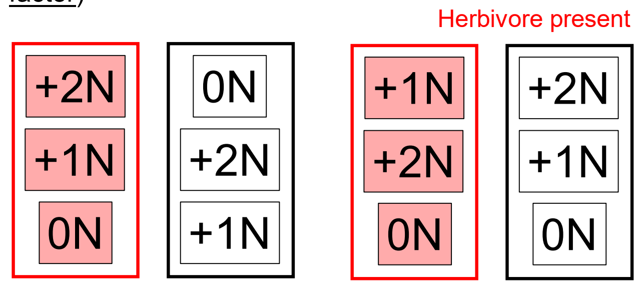
Confounding factors all over again!

	0N	**+1N**
Herb. absent	5	10
Herb. present	10	5

57

Split-plot design

- All N treatment levels (i.e., subplot factor) are represented in each block, which is itself a different herbivore treatment (i.e., whole-plot factor)



58

Questions?

59

Getting away from ANOVAs

60

Getting away from ANOVAs

- Thus far, I have presented IVs as categorical factors w/ different levels
- Indeed, ANOVA was invented by R.A. Fisher in the context of agricultural experiments
- But many times, different levels w/in a factor can be converted to continuous data

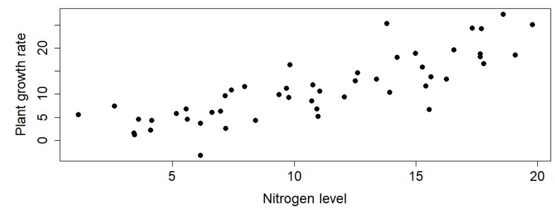
0N | +1N | +2N | +3N | … | +20N

0H | 1H | 2H | 3H | 4H | … | 20H

61

Regression

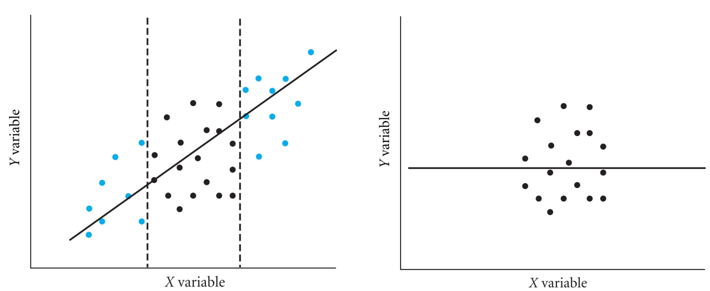
- Becomes regression instead of ANOVA
  - Can get change in DV ~ change in IV (slope)
  - Do predictions
  - Get goodness of fit (R^2)
- Instead of 10 replicates per level, focus on number of data points per IV



62

Regression

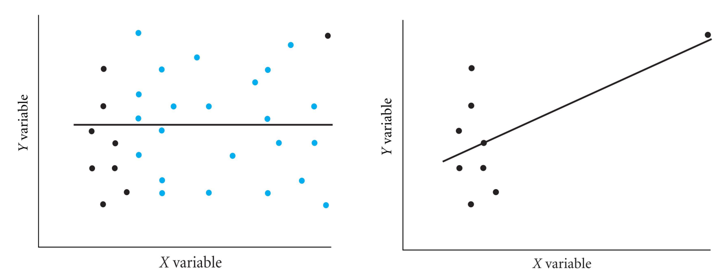
- Make sure IV is sampled over large enough range to get full range of response in DV



63

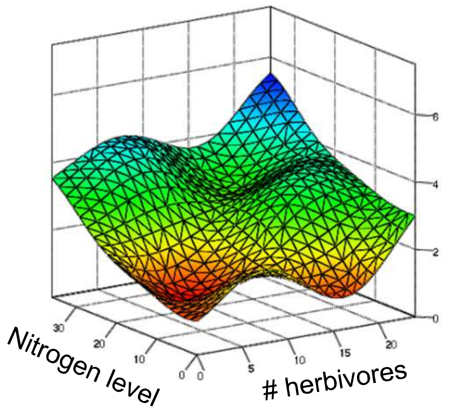
Regression

- Make sure IV is sampled uniformly within its range



64

Can get a response surface



65

Questions?

66

Summary

- Experiments account for sources of uncontrolled variation in DV → more powerful tests
- Accounts for confounding factors, non-independence, & noise in DV w/ replication, randomization, spacing btw plots, & blocking
- Try to make IV continuous rather than categorical
- In the end, it's all about thinking hard to account for confounding factors in your own research!

67

68

Statistics vignette

Brain cancer death rates in the US (example from Ellenberg, 2014)

Top 5 States
- South Dakota
- Nebraska
- Alaska
- Delaware
- Maine

Bottom 5 States
- Wyoming
- Vermont
- North Dakota
- Hawaii
- District of Columbia

69

Statistics vignette

Brain cancer death rates in the US (e

What's going on?
What do these ten states have in common?

Top 5
- Sou
- Nebraska
- Alaska
- Delaware
- Maine

- Vermont
- North Dakota
- Hawaii
- District of Columbia

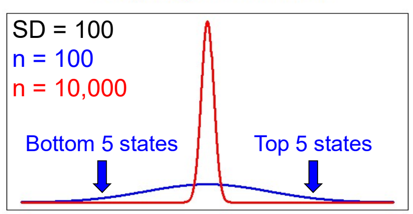
70

Small population sizes!

- Leads to noisy death rate estimates
  (related to Law of Large Numbers & standard errors)

$$SE = \frac{SD}{\sqrt{n}}$$

Sampling distribution



Brain cancer death rate

71