Are leprechauns the answer?

Quantitative Thinking in the Life Sciences

Nov 7th – Parsimony and using your textbook

Today

- Parsimony
- Using your textbooks
- Preparation for class project

Housekeeping

- No class next two weeks!
- Next class session is Nov 28th (full moon!)
- After today, only two class sessions left
- Homework C is due today
  - Chapter 8 R code: Modeling Elk populations in Rocky Mountain National Park
- This week's Plant and Soil Science Departmental talks 3 pm on Friday in 127 Jeffords
  - Meghan will talk worms!
  - I will be talk about the effects of global climate change on pests

A little jargon from Wiki

Independent variable

- predictor variable

Dependent variable

- response variable

X

Y

The principle of parsimony (AKA, Occam’s or Ockham’s Razor)

- Given competing hypotheses / models, a simpler explanation / model is better than a more complex one – if they both explain the same amount of information
- Why?

R Code!

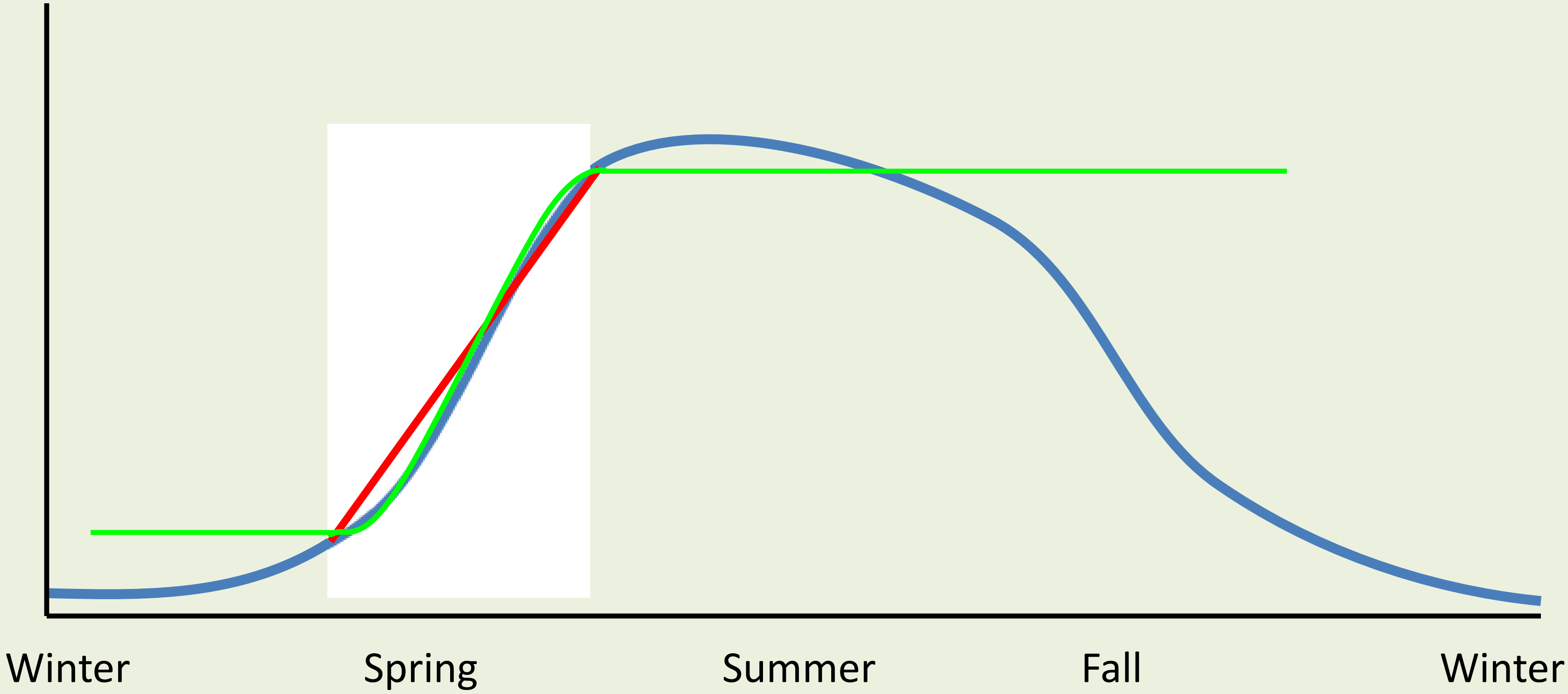
Possible explanations can get needlessly complex. It is coherent, for instance, to add the involvement of Leprechauns to any explanation, but Occam's razor would prevent such additions, unless they were necessary.

AIC and other information criterion

- Used to select the best model (e.g., model with the highest likelihood of being the best model)
  - Use the amount of error explained
  - Penalize models for use of extra fitted terms or extra parameters
  - R code:
    - > aic(model1)
    - > aic(model2)

Making choices

Banks grass mite populations



Page 7 of our book

Eight steps to successful data analysis

1. Decide what you are interested in
2. Formulate a hypothesis or hypotheses
3. Design the experiment or sampling routine
4. Collect dummy data. Make up approximate values based on what you would expect
5. Use the key here to decide on the appropriate test or tests
6. Carry out the tests using the dummy data
7. If there are problems go back to step 3 (or 2), otherwise collect the real data
8. Carry out the test(s) using the real data

Eight steps to successful data analysis

1. Decide what you are interested in
2. Formulate a hypothesis or hypotheses
3. Design the experiment or sampling routine
4. Collect dummy data. Make up approximate values based on what you would expect
5. Use the key here to decide on the appropriate test or tests
6. Carry out the tests using the dummy data
7. If there are problems go back to step 3 (or 2), otherwise collect the real data
8. Carry out the test(s) using the real data

Concept map

Data simulation

Today

Homework

Today

- Groups of 3ish
- Each of you will:
  - Reintroduce your system
    - Components – type of data collected
    - Relationships between components
  - As a group, go through the key (Pg 8) in the book and figure out an appropriate test
    - note the code
    - bookmark the page
- Extra time? – discuss the R code for your test(s). Does it make sense?
- Assignment D

Assignment D

R code only! No write-up

100 points. Due on November 28th

- Create an array/matrix with simulated data for your system
  - Column one should be your y variable (aka dependent or response variable)
    - TSS remaining, phosphorus, Resilience to Climate Change Index
  - The number of rows should be the number of expected samples
  - Columns 2-?: These are your x variables (aka independent or predictor variables)
    - year, site, pollen amount, harvest type, yield
- Populate the array with dummy data using the distribution and distribution functions from your homework (you don't need to put in measurement error right now)
- Using the test agreed upon today, write the code and test your hypothesis
- Send well annotated code to me as a single document (plus data files if you choose to use your real data)

Break into groups!

- Groups of 3ish
- Each of you will:
  - Reintroduce your system
    - Components – type of data collected
    - Relationships between components
  - As a group, go through the key (Pg 8) in the book and figure out an appropriate test
    - note the code
    - bookmark the page
- Extra time? – discuss the R code for your test(s). Does it make sense?
- Assignment D