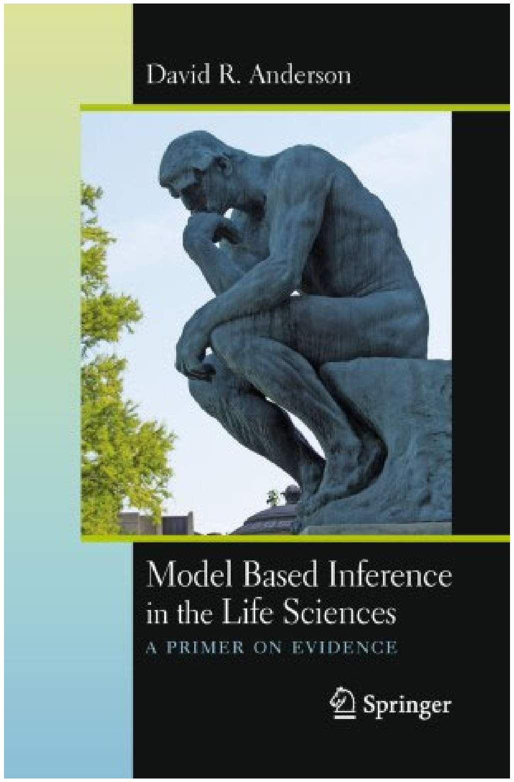
Anonymous peer reviews

- Due by midnight Sunday
- In journal article peer review format (no length requirement):
  - Summary
  - Major comments
  - Minor comments
  - (try not to use tracked changes)
- Constructive praise and criticism! Offer solutions if you have them, and if you're unsure you can evaluate something, say so.
- Be courteous, and evaluate what it is, not what you think it should be
- Check the final project rubric to understand which are key components
- More peer review resources on eLC

This week

- Discuss the concept of model-based inference
- Introduce the information theoretic approach to statistics
- Relating the scientific approach of alternative hypotheses to the statistical method of model selection and multimodel inference
- Practice formulating multiple working hypotheses and turning it into multiple models
- Learn how to do model selection responsibly
- Put approaches in the context of modeling goals



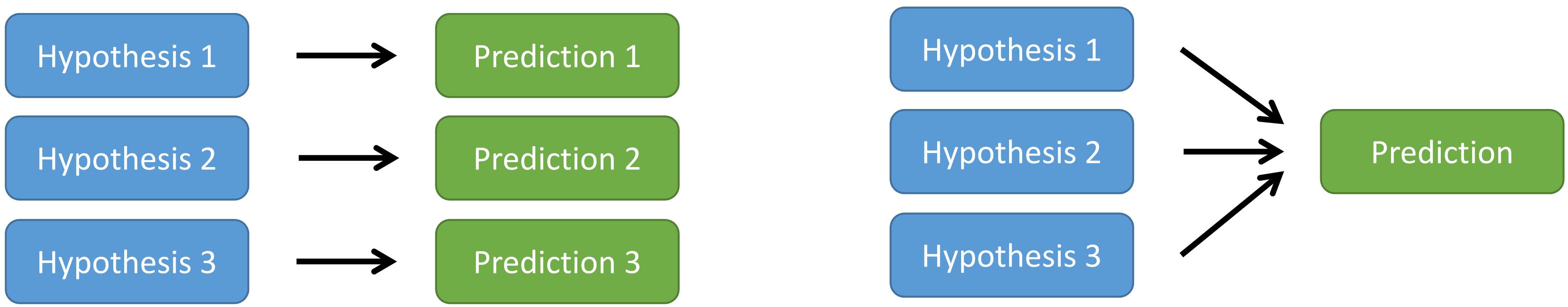
Before we talk about model selection, and now that we’re at the end of the semester…what is a model?

What is point of having models in science?

Models can represent scientific hypotheses

What is a scientific hypothesis?

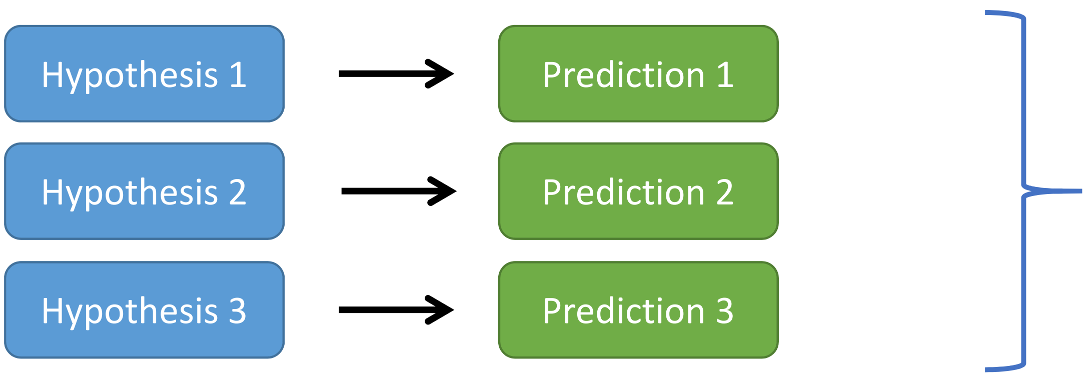
What makes a "good" scientific hypothesis?



Models can represent scientific hypotheses

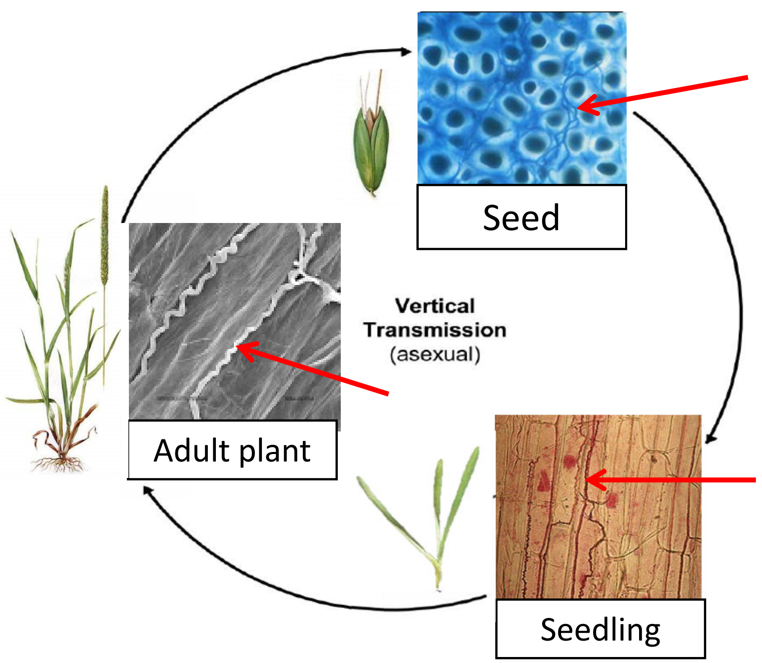
What is a scientific hypothesis?

What makes a “good” scientific hypothesis?



We may have a set of *alternative hypotheses* that correspond to a single scientific question
These are represented by *alternative models*

For example…
How does an *Epichloë* endophyte alter the vital rates of *Poa alsodes*?



Goal: Find best model that describes effects of endophyte on growth, survival, and reproduction

Alternative hypotheses (for each vital rate response):

- H1: Endophytes have no effect
- H2: Endophyte has an overall (main) effect only
- H3: Endophyte has no overall (main) effect but only an effect depending on the size of the plant (interaction)
- H4: Endophyte has both overall effect and a size-dependent effect (main effect and interaction)

For example…
How does an *Epichloë* endophyte alter the vital rates of *Poa alsodes*?

Alternative hypotheses (for each vital rate response):

- H1: Endophytes have no effect
- H2: Endophyte has an overall (main) effect only
- H3: Endophyte has no overall (main) effect but only an effect depending on the size of the plant (interaction)
- H4: Endophyte has both overall effect and a size-dependent effect (main effect and interaction)

Using survival as an example, alternative models:

For example…
How does an *Epichloë* endophyte alter the vital rates of *Poa alsodes*?

Alternative hypotheses (for each vital rate response):
- H1: Endophytes have no effect
- H2: Endophyte has an overall (main) effect only
- H3: Endophyte has no overall (main) effect but only an effect depending on the size of the plant (interaction)
- H4: Endophyte has both overall effect and a size-dependent effect (main effect and interaction)

Using survival as an example, alternative models:
- H1: Surv(t+1) ~ Size(t)
- H2: Surv(t+1) ~ Endo + Size(t)
- H3: Surv(t+1) ~ Size(t) + Endo:Size(t)
- H4: Surv(t+1) ~ Endo + Size(t) + Endo:Size(t)
  (the same as Surv(t+1) ~ Endo*Size(t) in R)

Practice

- Think of one of your research questions
- Take 10 mins and come up with a set of alternative hypotheses
- Break-out groups (15 mins):
  - Discuss your set of alternative hypotheses with each other
  - What makes a good set of alternative hypotheses? What did you find easy or difficult in coming up with your set of alternative hypotheses?
  - Report out

What makes an effective set of alternative hypotheses?

The idea of model selection

“A model is a simplification or approximation of reality and hence will not reflect all of reality. … Box noted that “all models are wrong, but some are useful.” While a model can never be “truth,” a model might be ranked from very useful, to useful, to somewhat useful to, finally, essentially useless.”

Burnham and Anderson (2002)

Information theory: a general philosophy for evaluating alternative hypotheses in the face of evidence

“Kullback-Leibler information” is a function that denotes:

1. The amount of “information” lost when a model is used to approximate full reality
2. The “distance” from the model to full reality

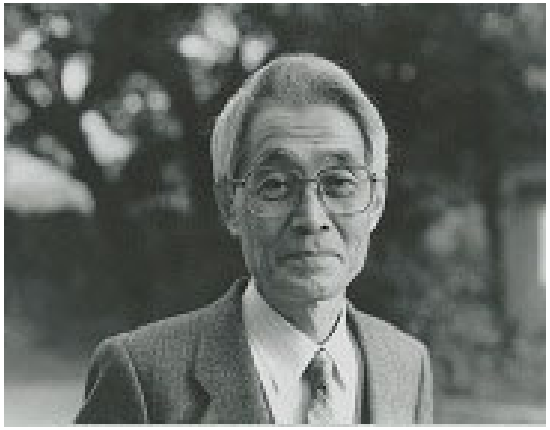
The “best” model of a set of alternative models would be the one that has the smallest amount of information loss, or the shortest distance away from full reality

Problem: We usually can’t measure “full reality”

The Akaike Information Criterion helps solve this issue by linking information theory with statistical principles

As of October 2014, the 1974 paper has received more than 14,000 citations in the Web of Science: making it the 73rd most-cited research paper of all time (Van Noorden et al. 2014).

Akaike, H. (1974), "A new look at the statistical model identification" IEEE Transactions on Automatic Control 19 (6): 716–723 doi:10.1109/TAC.1974.1100705, MR 0423716.



Hirotugu Akaike

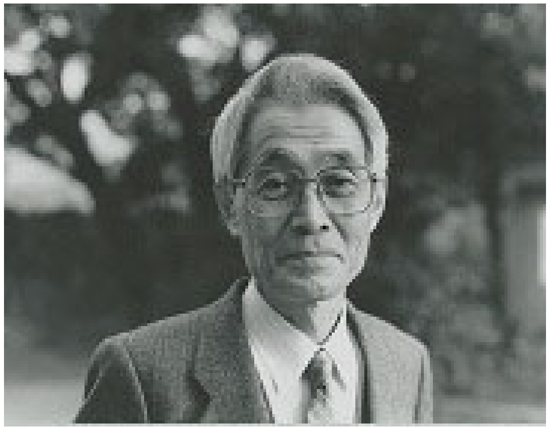
The Akaike Information Criterion helps solve this issue by linking information theory with statistical principles

Kullback-Leibler information =
log likelihood of parameters given the data – number of estimable parameters =

$$\log(L_n(\hat{\theta}|data)) - K$$

What is likelihood?
- The probability of the observed data given the parameter values of the model.
- The specific likelihood function is defined differently depending on the probability distribution.
- The important thing is that this is something we can calculate!

This subtraction corrects for bias



Hirotugu Akaike

The Akaike Information Criterion helps solve this issue by linking information theory with statistical principles

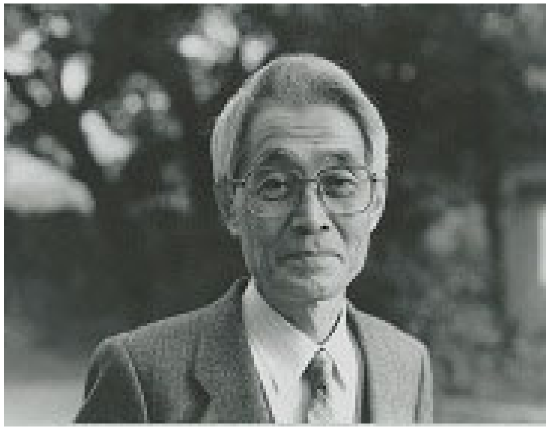
Kullback-Leibler information =
log likelihood of parameters given the data – number of estimable parameters =

$$\log(L_n(\hat{\theta}|data) - K$$

To make the math involved be more harmonious with other stuff we'd like to do with this estimate, Akaike multiplied it all by -2, and thus

Akaike's Information Criterion (AIC) =

$$-2\log(L_n(\hat{\theta}|data) + 2K$$



Hirotugu Akaike

For a set of models, the bigger the AIC value, the farther away the model is from full reality

The Akaike Information Criterion helps solve this issue by linking information theory with statistical principles

For a set of models, the bigger the AIC value, the farther away the model is from full reality

$$\mathrm{AIC} = -2\log\left(L_n\left(\hat{\theta}\middle|data\right)\right) + 2K$$

- This value is also known as "deviance"
- Measures how badly a model fits (the bigger the value the worse the fit)
- Becomes smaller the more parameters added to model (fit becomes better)
- But adding more estimable terms adds uncertainty to model

This bias correction is based on the number of estimable parameters *K*, which also can be interpreted to mean a correction/penalty for the uncertainty due to additional parameters

Models with *better fit using a smaller number of parameters (parsimony)* will result in a smaller AIC value, thus interpreted to have smallest information loss and be closest to full reality

For small sample sizes, use corrected AIC (AICc)

$$\mathrm{AICc} = \underbrace{-2\log\left(L_n\left(\hat{\theta}\middle|data\right)\right)}_{\text{Deviance, measures lack of fit}} + \underbrace{2K}_{\text{Correction/penalty for uncertainty due to additional parameters}} + \underbrace{\frac{2K(K+1)}{n-K-1}}_{\text{Correction for small sample size (n/K <40)}}$$

Where *n* is the sample size (total number of data points) and *K* is the number of estimable parameters

For example…
How does an *Epichloë* endophyte alter the vital rates of *Poa alsodes*?

Alternative hypotheses (for each vital rate response):

- H1: Endophytes have no effect
- H2: Endophyte has an overall (main) effect only
- H3: Endophyte has no overall (main) effect but only an effect depending on the size of the plant (interaction)
- H4: Endophyte has both overall effect and a size-dependent effect (main effect and interaction)

Using survival as an example, alternative models:

- H1: Surv(t+1) ~ Size(t)
- H2: Surv(t+1) ~ Endo + Size(t)
- H3: Surv(t+1) ~ Size(t) + Endo:Size(t)
- H4: Surv(t+1) ~ Endo + Size(t) + Endo:Size(t)
  (the same as Surv(t+1) ~ Endo*Size(t) in R)

AIC gives us a way to directly compare among alternative models and pick the most parsimonious explanation!