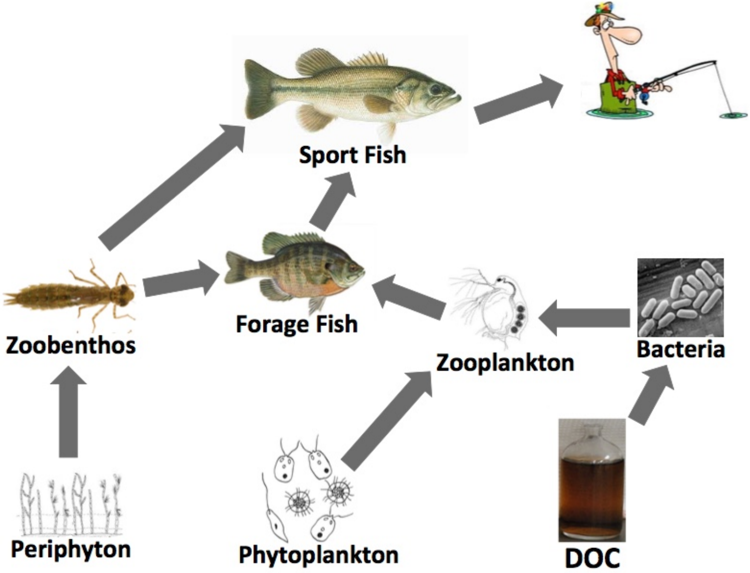
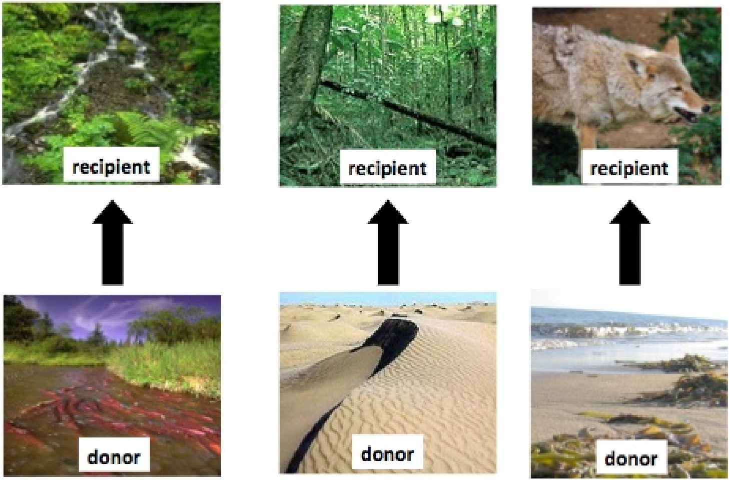
Lecture 16 - Statistical applications: Kelly *et al.* 2014 & Maximum likelihood

A little background on lake foodwebs. . .



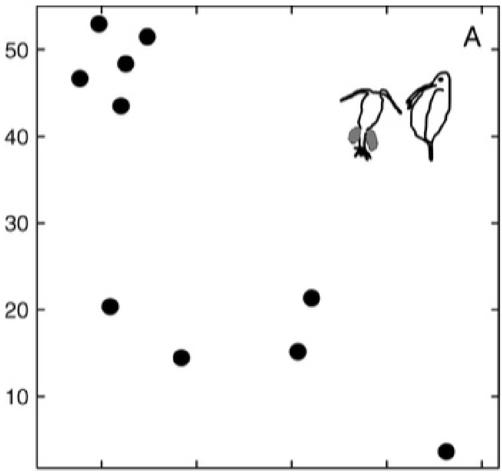
What's a resource subsidy?

What's a resource subsidy?



How did Kelly *et al.* ask whether terrestrial DOC was a subsidy for lakes?

How did Kelly *et al*. ask whether terrestrial DOC was a subsidy for lakes?



How did Kelly *et al.* ask whether terrestrial DOC was a subsidy for lakes?

Table 1. Summary of water chemistry and planktivore characteristics for each survey lake.

Lake	Area (ha)	Temperature (°C)	DOC (mg/L)	K_d (m^{-1})	Chlorophyll *a* (µg/L)	Total N (µg/L)	Total P (µg/L)	*Chaoborus* density (no./ m^2)	Mixed layer depth (m)	Piscivore present/ absent
Bay	67.30	22.7 (1.1)	5.9 (0.7)	0.99	4.5 (0.7)	427.8 (5.6)	14.6 (3.8)	154.4 (47.4)	2.7 (0.1)	present
Bergner	17.85	22.1 (0.9)	11.8 (2.5)	1.50	5.9 (0.8)	497.7 (13.5)	21.5 (4.7)	430.2 (113.5)	3.3 (0.3)	present
Brown	32.57	21.8 (1.0)	9.3 (0.6)	1.09	8.9 (3.7)	511.1 (27.0)	36.3 (8.6)	58.0 (29.4)	3.0 (0.0)	present
Crampton	25.81	21.9 (1.0)	5.4 (1.1)	0.77	5.6 (0.7)	430.6 (52.9)	20.1 (5.4)	156.5 (47.7)	5.0 (0.2)	present
Hummingbird	0.76	23.3 (1.3)	25.9 (1.6)	3.09	14.2 (4.0)	903.8 (88.7)	35.0 (7.9)	1113.7 (305.2)	0.8 (0.1)	present
Inkpot	6.61	22.3 (0.9)	11.4 (2.2)	1.23	3.9 (0.5)	451.1 (13.8)	25.8 (5.7)	232.4 (81.6)	2.7 (0.1)	present
Long	7.87	21.6 (1.1)	8.0 (0.4)	1.83	8.2 (1.8)	422.1 (7.91)	11.3 (1.5)	955.4 (180.2)	1.3 (0.1)	present
Morris	5.93	22.8 (1.1)	17.4 (2.1)	3.22	9.0 (3.0)	709.4 (72.6)	33.3 (7.6)	405.9 (88.7)	1.1 (0.1)	present
Raspberry	4.63	23.2 (1.0)	6.4 (0.4)	1.28	5.0 (0.8)	486.1 (29.7)	27.4 (6.2)	334.6 (133.4)	1.7 (0.1)	present
Reddington	1.24	24.2 (1.1)	22.0 (1.3)	4.64	9.4 (2.3)	693.5 (115.2)	33.8 (8.7)	178.8 (36.3)	0.9 (0.0)	absent

Notes: Values are means with SE in parentheses. DOC stands for dissolved organic carbon. The diffuse light attenuation coefficient is K_d.

How did Kelly *et al.* ask whether terrestrial DOC was a subsidy for lakes?

TABLE 2. Model comparisons for relationships between zooplankton production and dissolved organic carbon (DOC), light attenuation (K_d), total phosphorous (TP), chlorophyll *a*, piscivore presence/absence (piscivore), and *Chaoborus* density (Chaob).

Model	Hypothesis	Slope	r^2	P
Production ~ K_d	t-OC	−11.31	0.60	<0.01
Production ~ DOC	t-OC	−1.76	0.46	0.03
Production ~ chlorophyll *a*	resource	−1.12	0.37	0.06
Production ~ piscivore	predation	31.23	0.28	0.11
Production ~ TP	resource	−0.99	0.23	0.16
Production ~ Chaob	predation	−0.02	0.04	0.69

Note: Models were categorized based on hypothesized drivers of zooplankton production (terrestrial organic carbon [t-OC] concentration, resource availability, or predation).

How did Kelly *et al.* ask whether terrestrial DOC was a subsidy for lakes?

We evaluated the support for different hypothesized drivers of zooplankton production, including t-OC, lake productivity, and predation by planktivores, by comparing simple regression models. Direct quantification

Stuart's version of statistics

My (short) version of statistics relies on two concepts:

- random variables
- linear, and sometimes simple non-linear, models

Random variables

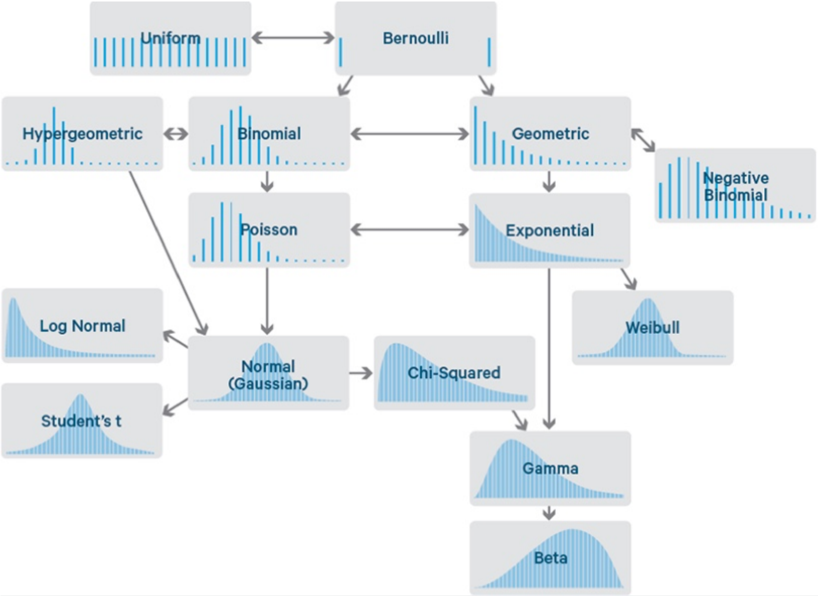
- a way to quantify the outcomes of a random process
- a way to map a random process to numbers

Repeated quantification of the same random process begets a probability distribution!

Probability distributions

- describe a variety of random processes and therefore have a number of different characteristics and shapes
- can be used to test hypotheses or models

Probability distributions



Probability distributions

- describe a variety of random processes and therefore have a number of different characteristics and shapes
- can be used to test hypotheses or models
- we can do this using something called a **Likelihood**

```
L(model|data) = P(data|model)
```

Example with coin flips

Imagine we flipped a coin twice and got two heads. . .

What are a couple hypothesis or models for our coin?

Example with coin flips

What is the likelihood of our two models?

Example with coin flips

What is the likelihood of our two models?

What if we flipped 8 more times and got 3 heads and 5 tails?

What is the likelihood of our two models?

Linear models

- familiar with the equation for a line?
- that is a linear model, but isn't exactly what is meant by a "linear model"...
- how did Kelly *et al.* use linear models?

linear models + probability distributions = biological hypothesis testing

our coin flip example didn't include a linear model, but it did have a probability distribution

so far our discussion of Kelly *et al.* 2014 hasn't had a probability distribution, but how might we add one?

what does adding this probability distribution accomplish?

what does adding this probability distribution accomplish?

- accounts for "variation"
  - measurement error
  - genetic variation across populations
  - variation in growth rate across individuals
  - etc.
- allows us to compare one model to another, including a null (purely random) model
  - do this using Likelihoods

Wednesday

talk about how to translate our models into code and confront with data

practice with some data from Kelly *et al*. 2014 and other examples