Welcome to Ecological Statistics and Data! (Bio 133)

Instructor: Eric Scott

Why “Ecological” statistics and data?

- Ecologists often count things and watch when they reproduce and die
- These types of data have special properties
- Violate assumptions of statistical tests typically covered in more general or introductory courses
- Rather than trying to force data to match one of a list of statistical tests, use flexible tools to model data generated by a variety of ecological processes

Introductions

The team:

- Eric Scott (eric.scott@tufts.edu)
- Avalon Owens (avalon.owens@tufts.edu)

The class:

- Undergraduate and graduate students from several departments with varied backgrounds in math, statistics, computer programming, and ecology.

Today's Outline

1. Logistics
2. Probability
3. Homework

Logistics

Take a look at the syllabus:

- General format (lectures mixed with computer exercises – find a partner or 2)
- Books: Bolker, R4DS
- Access to laptops
- Lab

Textbook Definition of Probabilty

"If an observation is made *N* times and event *A* occurs N_A times, then with a high degree of certainty, the relative frequency of N_A/N is close to *P(A)*, the probability of *A* in a single trial, $P(A) \approx N_A/N$, provided *N* is sufficiently large."

- "event" = something that does or does not occur
- "trial" = single observation or data point or "experiment" in which the event can occur or not

- P(A) = probability that event (A) will occur in a single trial
- N = total number of trials
- N_A = number of times A occurs

$$P(A) = \lim_{N \to \infty} \left(\frac{N_A}{N} \right)$$

Example application: Demography of a perennial wildflower

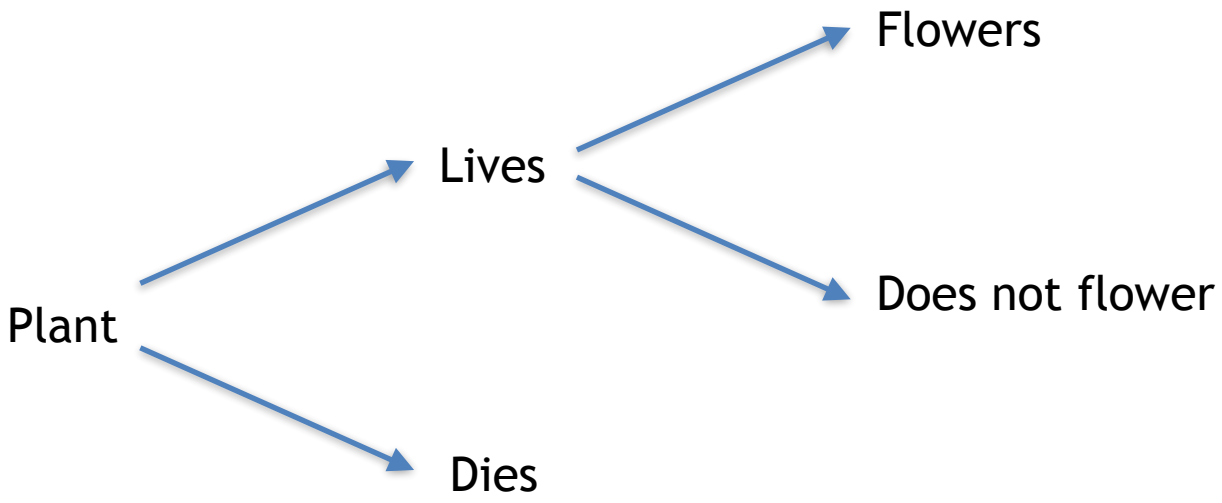
Discovering what N is for your homework

$$P(A) = \lim_{N \to \infty} \left(\frac{N_A}{N} \right)$$

P(survival) = 4/5 = 0.8

P(flowering) = 1/4 = 0.25

Plant #	fate
1	Flowers
2	Vegetative
3	Vegetative
4	Vegetative
5	Dead

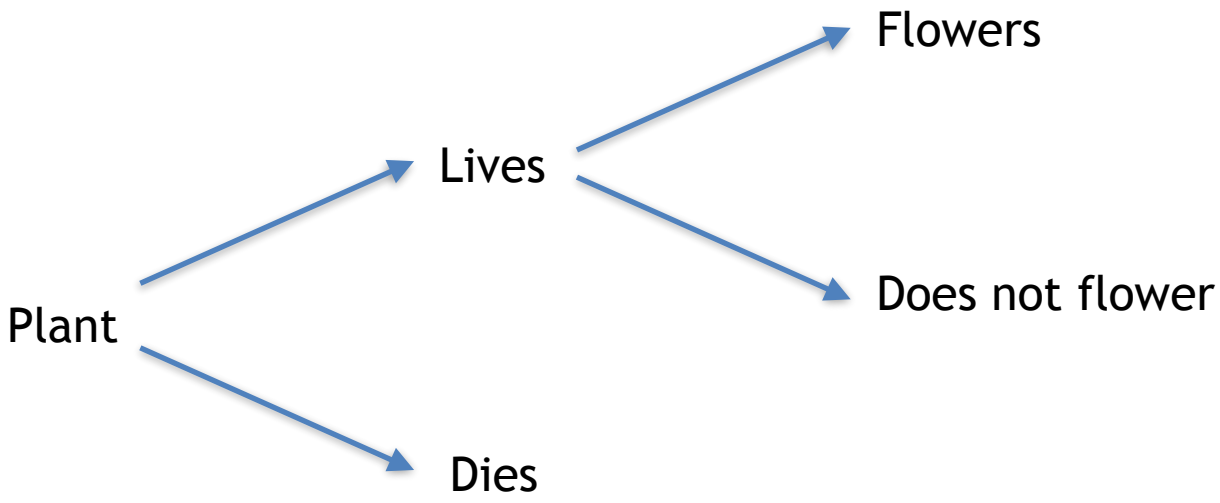


Example application: Demography of a perennial wildflower

What is the probability that a plant survives AND flowers?

Need to know about the rules of probability…

Plant #	fate
1	Flowers
2	Vegetative
3	Vegetative
4	Vegetative
5	Dead



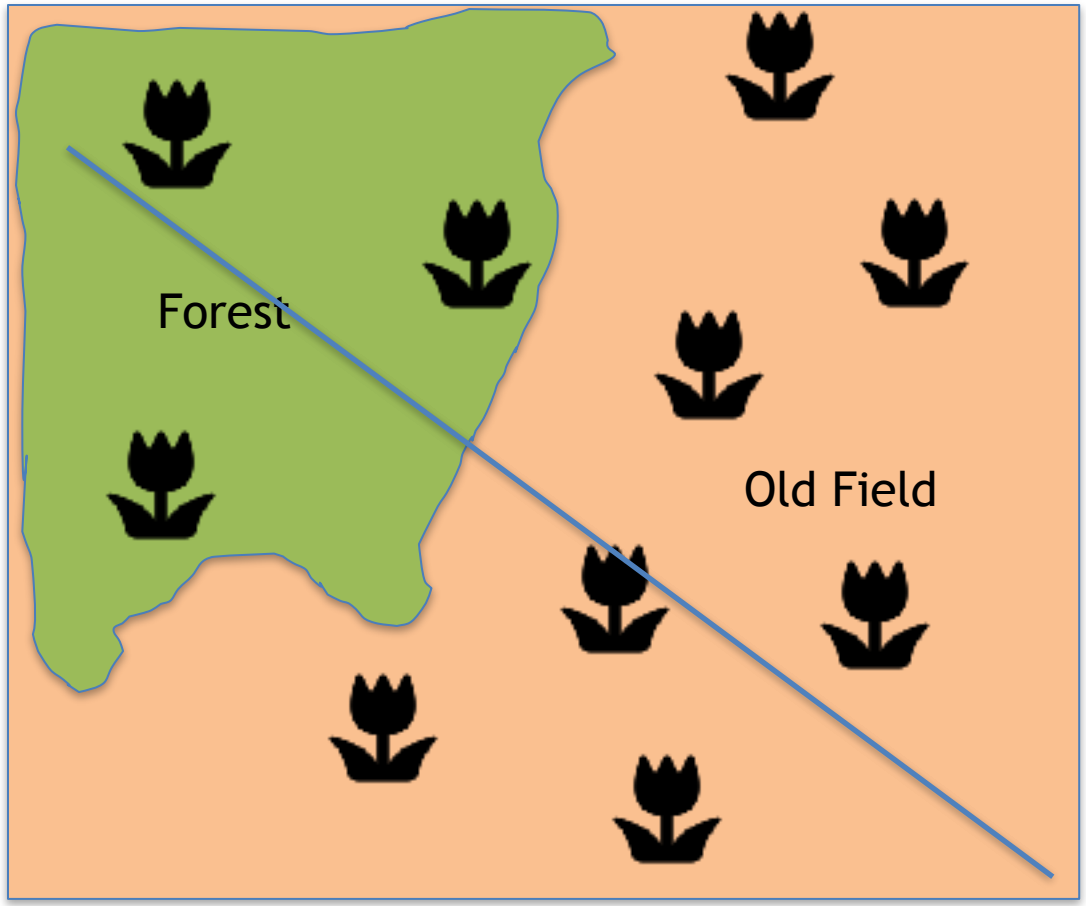
Homework

- Measure a probability of some event by sampling a population.

Sampling and Scope of Inference

- Definition of probability assumes trials are representative of a "population"
- "population" (statistics definition) = larger group for which you estimating the probability
- Trials should be **representative** and **independent**.
- Scope of inference = "population" over which your data are **representative** and **independent** observations

Types of Sampling



- Random
  - Label all plants in population with numbers. Choose subset at random
- Systematic
  - Walk a transect and sample individuals every 10 paces
- Haphazard
  - Choose samples without any conscious bias (Be careful! Unconscious bias exists!!!)
- Stratified
  - Sample in forest and old field separately. # samples in each determined by relative area and variability