NRES_798_2_201501

Review of probability, distributions, statistical inference

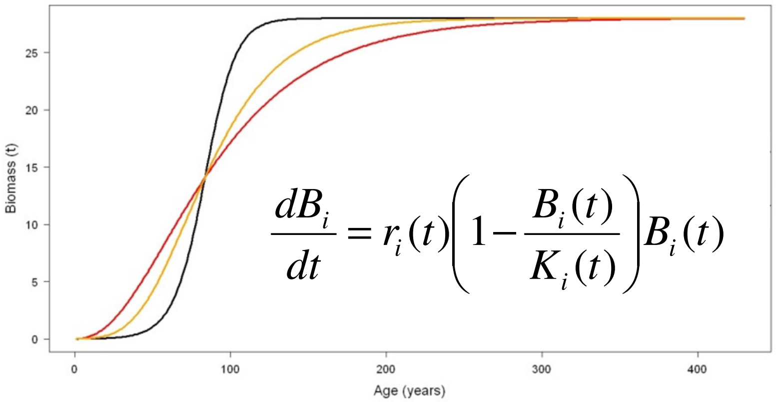
Interest

Area	Interest
ANOVA (ANCOVA)	4
Spatial statistics	4
Time series	2
Habitat modeling	1
Mixed models	1
Multivariate stats	1

Stats as tool to illuminate the interesting questions in biology (variability)

Stats as puzzle to be solved (multiple answers)

What is an scientific model?



$$\frac{dB_i}{dt} = r_i(t)\left(1 - \frac{B_i(t)}{K_i(t)}\right)B_i(t)$$

A model is a simplified version of reality that is developed to:

- Test a hypothesis
- Better understand how an ecological system functions
- Predict how a ecological system will change in response to future shifts

Statistics help us test and inform models

Statistics focuses on uncertainty in data, and models

What is a model?

model = simplified version of reality

Verbal model

Qualitative model

Quantitative model

Complex quantitative model

The persistence of marmot populations depends on winter temperatures and the availability of food

model = simplified version of reality

Verbal model

Qualitative model

Quantitative model

Complex quantitative model

Marmot populations increase when food is abundant.

X1: fed
X2: not fed

Ho: $N_{x1} = N_{x2}$
Ha: $N_{x1} > N_{x2}$

model = simplified version of reality

Verbal model

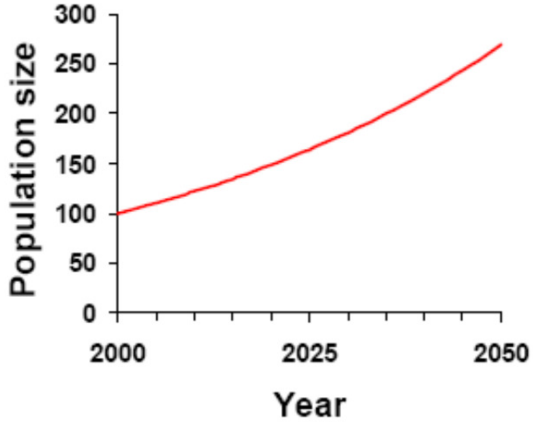
Qualitative model

Quantitative model

Complex quantitative model

$N_{t+1} = (b + s)*N_t$

s = overwinter survival = 0.50
b = birth rate = 0.32 + food*0.015 = 0.52



model = simplified version of reality

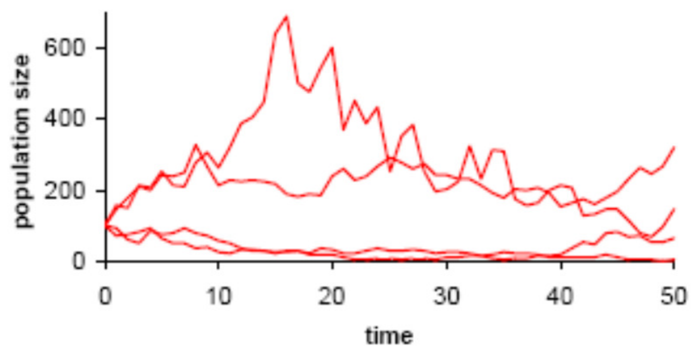
Verbal model

Qualitative model

Quantitative model

Complex quantitative model

$$N_{it} = \frac{k_{it} N_0 e^{r_{it} t}}{k_{it} + N_0 (e^{r_{it} t} - 1)}$$



What is a model?

model = simplified version of reality

Verbal model

Qualitative model

Quantitative model

Complex quantitative model

Marmot numbers increase when food is abundant.

X1: fed
X2: not fed

Ho: $N_{x1} = N_{x2}$
Ha: $N_{x1} > N_{x2}$

Why stats? Sample vs. Population

What is a model?

model = simplified version of reality

Verbal model

Qualitative model

Quantitative model

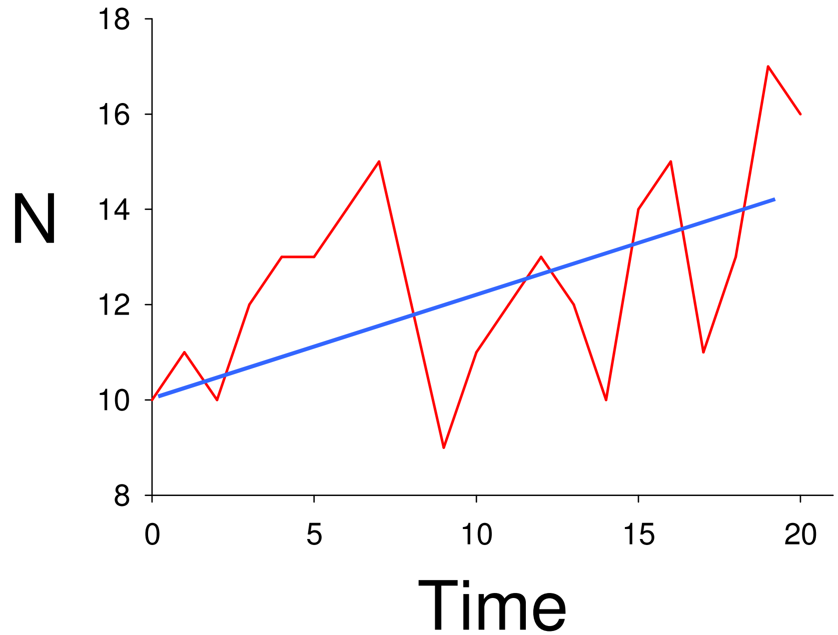
Complex quantitative model

- Assumptions explicitly stated
- Interactions between factors explicit
- Increased quantification

Increased partitioning of uncertainty

Population response

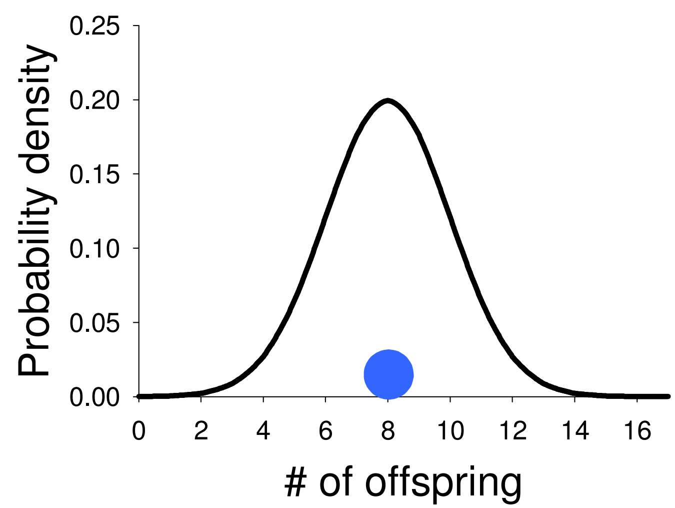
$$N_{t+1} = N_t + \text{births} + \text{immigration} - \text{deaths} - \text{emigration}$$



Real population

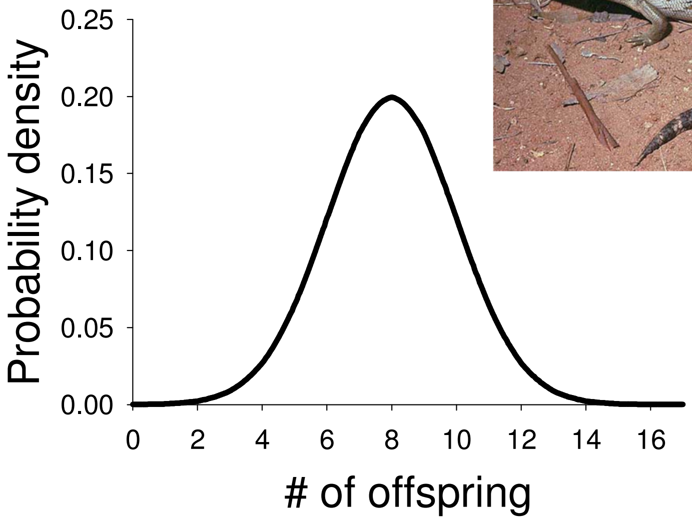
Estimated population

Blue tongue lizard

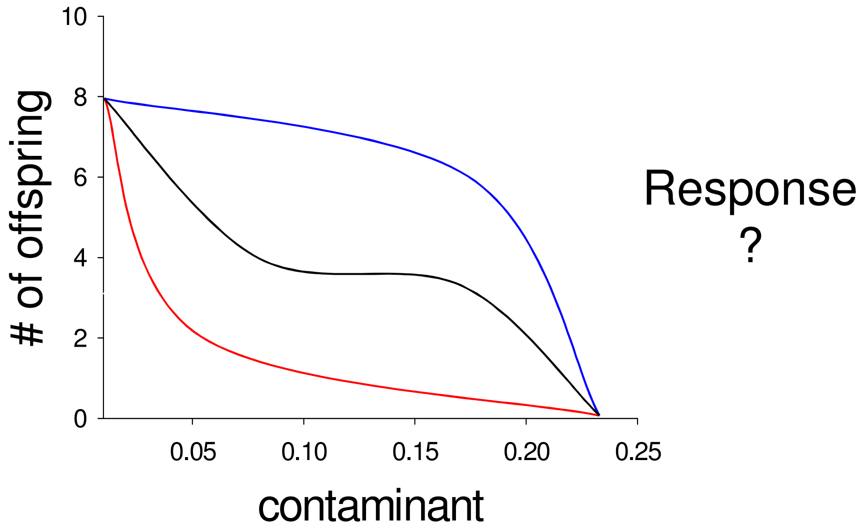


Uncertainty in ecology

- Variability
  - naturally occurring
  - attributable to ‘true’ heterogeneity in a population



- Incertitude
  - arises due to lack of knowledge about parameters or models
  - can be reduced by collecting more and better data



Uncertainty in ecology

Four types of uncertainty

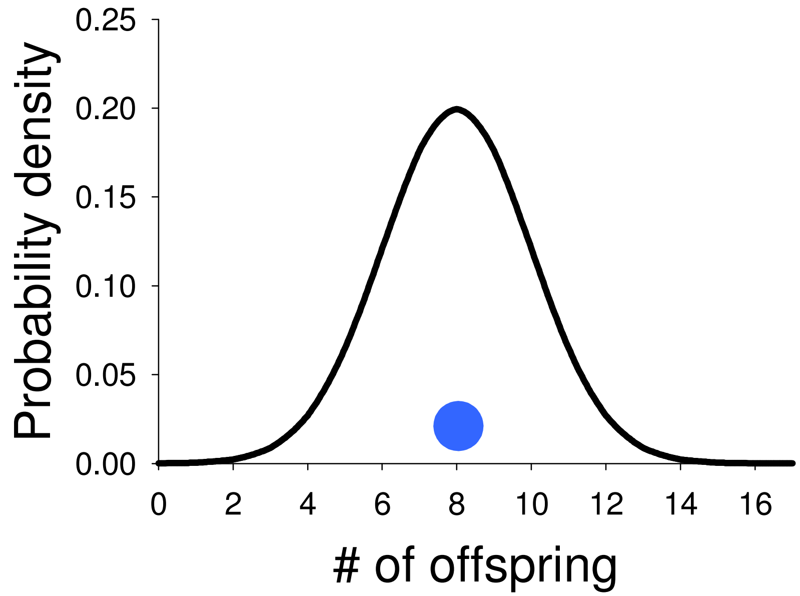
- Natural variation
- Measurement error
- Systematic uncertainty
- Model uncertainty

Incertitude !

Natural variation

Blue tongue lizard

- Genotype
- Phenotype
  (gene * environment)
- Behaviour
- Animal condition
- Environment

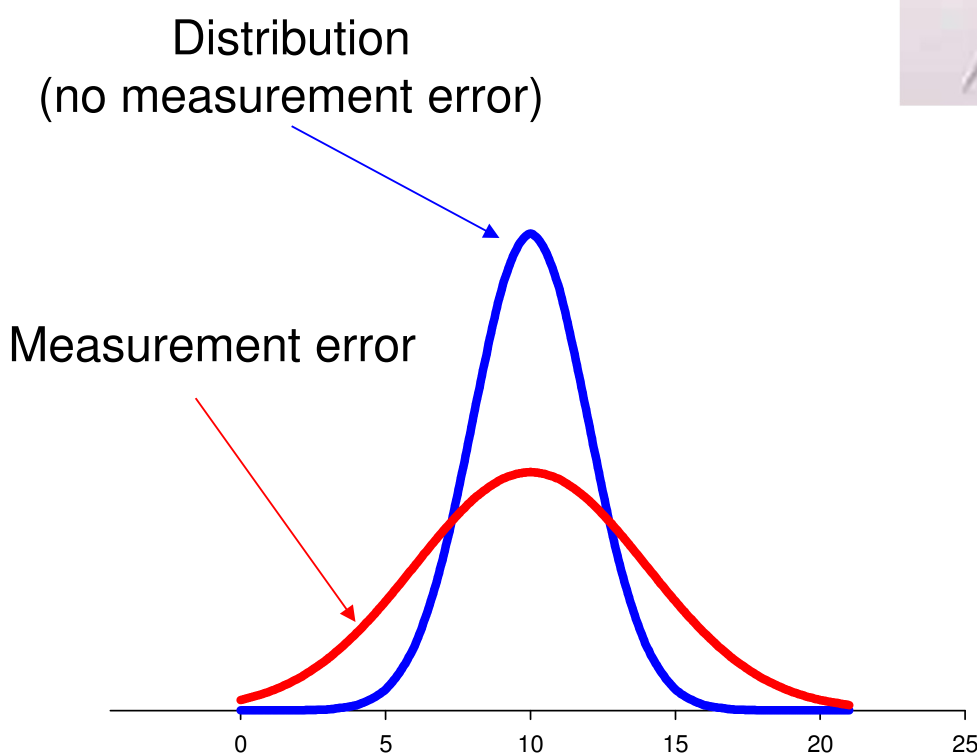


Uncertainty in ecology

- Measurement error

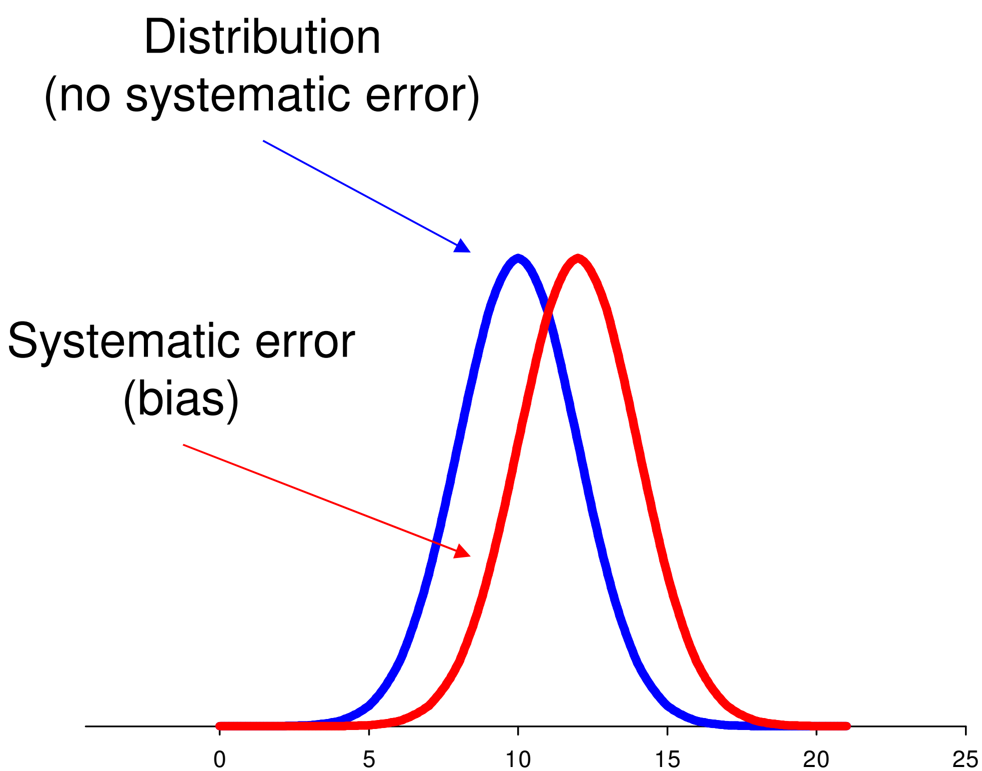
Blue whale

Beaked whale



Uncertainty in ecology

- Systematic error



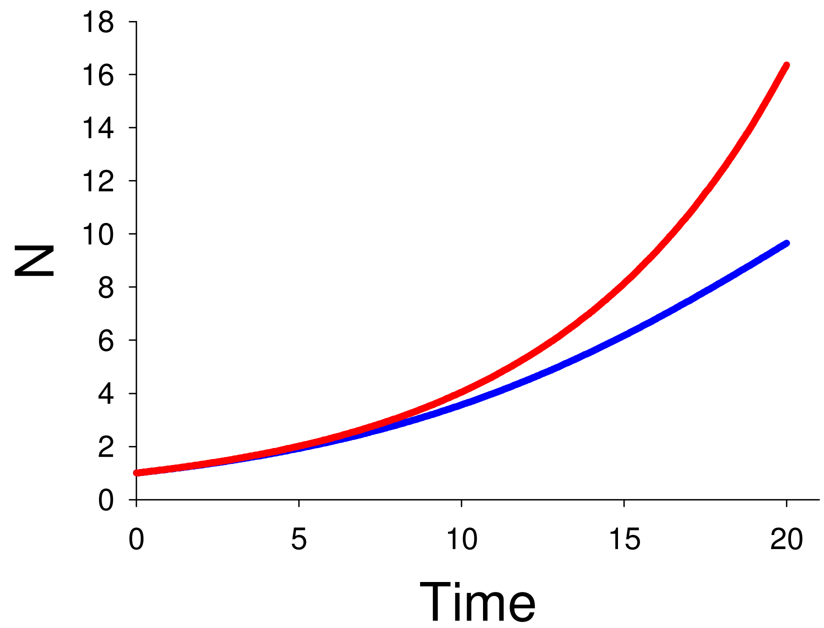
Uncertainty in ecology

- Model uncertainty: biological process

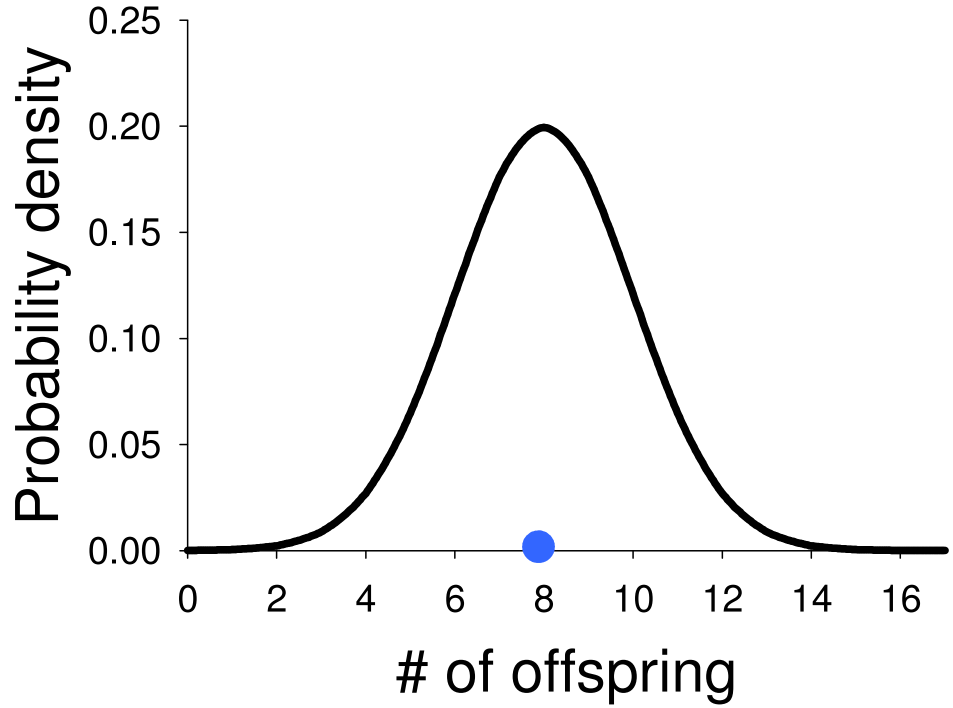
$$N_{t+1} = N_t + N_t * r$$

or

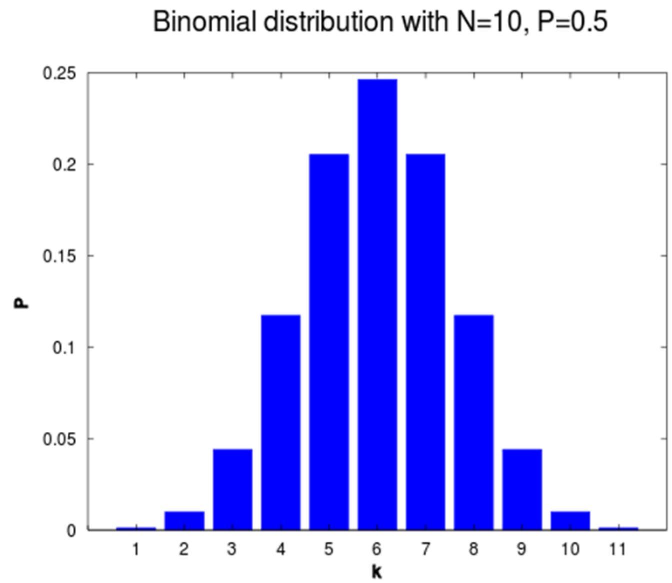
$$N_{t+1} = N_t + (N_t * r * (1 - N_t / k))$$



Blue tongue lizard



- Uncertainty, variation, stochasticity
  - Some times interesting
  - Some time a nuisance
- Creates world where statistics are necessary
  - Characterize
  - Test



Observed variation ~ probability distribution

Frequents and Bayesian statistics

Probability distribution

- Aggregation of observations of a **random variable**
- Probabilities of a single random variable taking on various alternative values

Random variable

- Function that assigns a numerical value to each possible outcome of an experiment
- Numerical set that is defined by the **probability** of given discrete outcomes of an experiment
  - E.g. Probability associate with having 0,1,2,3 offspring during a single reproductive event

Probability review

Event
- fair toss of neutral coin

Discreet outcome
- can be assigned a positive integer

Sample space
- head, tail
- **set** of all possible outcomes
  - Outcomes mutually exclusive
  - Outcomes in set exhaustive

Probability review

Estimating the probability of a discreet event (by sampling)

$$P = \frac{number\ of\ times\ for\ outcome}{number\ of\ samples}$$

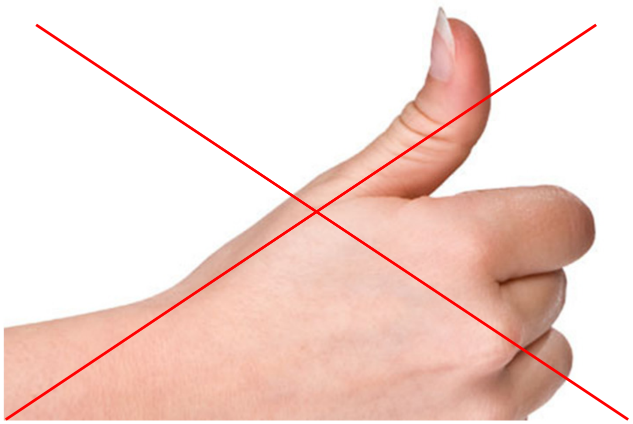
$$0.0 \leq P \leq 1.0$$

$$\sum_{i=1}^{n} P(A_i) = 1.0$$

First axiom of probability

Assumptions

- Fair toss



Initial state of system unknown, or ignored

Dynamics of system unknown, or ignored

Unable or unwilling to measure, “random”

- Probability estimate dependent on how the sample space is defined
  - Dependent on how our question structured
  - Dependent on the assumptions we are willing to make
  - Depends on the details of the trials
- How your study is defined is very important (need to be specific)
  - Clear, explicit hypothesis
  - What is the sample space that you are sampling
  - What variability are you interested in, what are you willing to treat as noise

Simple probabilities

- Whirligig beetle
  - Always produces exactly 2 litters (assumption?)
  - Either 2,3 or 4 offspring per litter

- Questions about whirligig fitness
  - What is the probability of producing N offspring?

Possible fitness = sample space =

Fitness = {(2,2),(2,3),(2,4),(3,2),(3,3),(3,4),(4,2),(4,3),(4,4)}

Fitness outcomes exhaustive
Probability of each outcome = 1/9

Possible fitness range 4 to 8 offspring

Combining simple probabilities

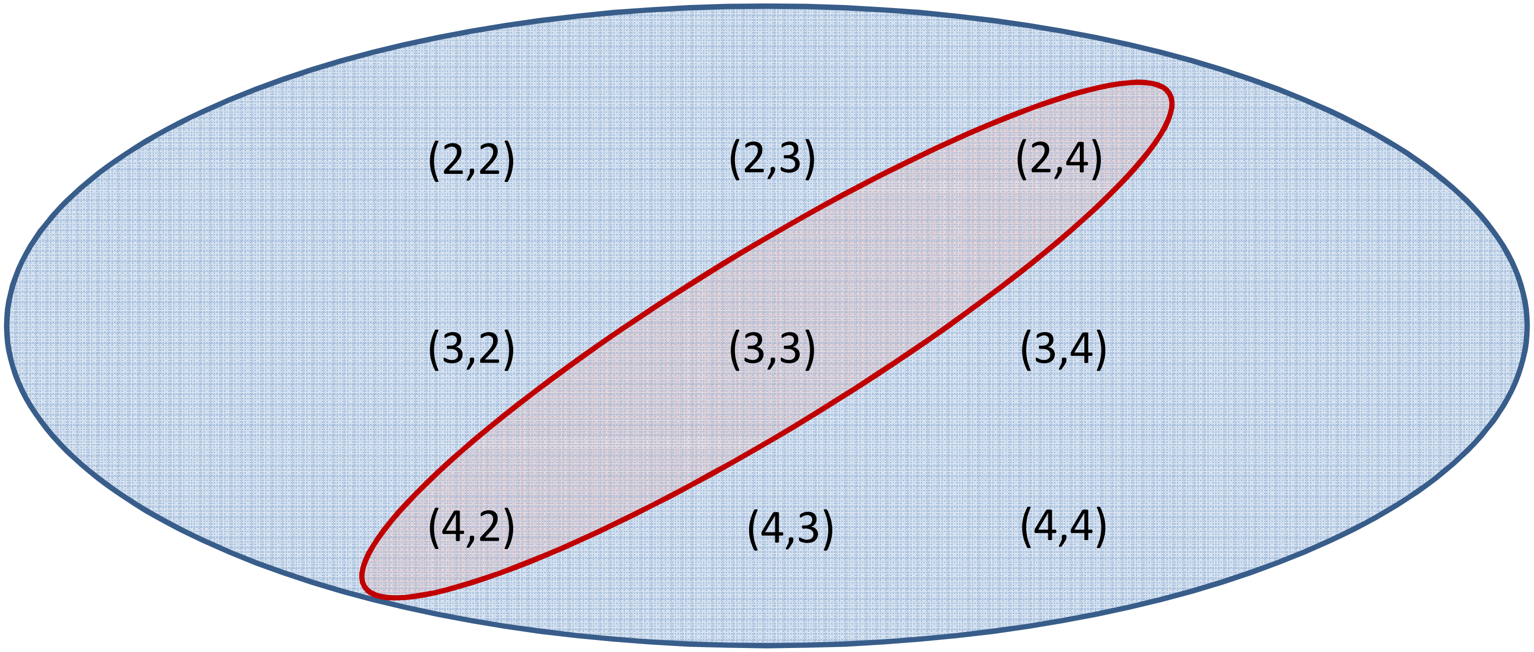
- What is the probability of observing a whirligig that has 6 offspring?

- 6 offspring = {(2,4),(3,3),4,2)}

- Complex event: composite of multiple simple probabilities

Combining simple probabilities

- What is the probability of observing a whirligig that has 6 offspring?



Fitness = {(2,2),(2,3),(2,4),(3,2),(3,3),(3,4),(4,2),(4,3),(4,4)}
6 offspring = {(2,4),(3,3),4,2)}

Complex events

Subset: 6 offspring subset of Fitness

Complex events

Fitness = {(2,2),(2,3),(2,4),(3,2),(3,3),(3,4),(4,2),(4,3),(4,4)}
6 offspring = {(2,4),(3,3),4,2)}

Assume # offspring in second litter is **independent** of first litter

Complex event probability (**union** of simple events)
P(6 offspring) = P((2,4) **or** (3,3) or (4,2) = P((2,4)) + P((3,3)) + P((4,2))

P(6 offspring) = P(A or B or C) = P(A) + P(B) + P(C)

The probabilities of a complex event equals the sum of the component simple events.

Second axiom of probability

Combining simple events

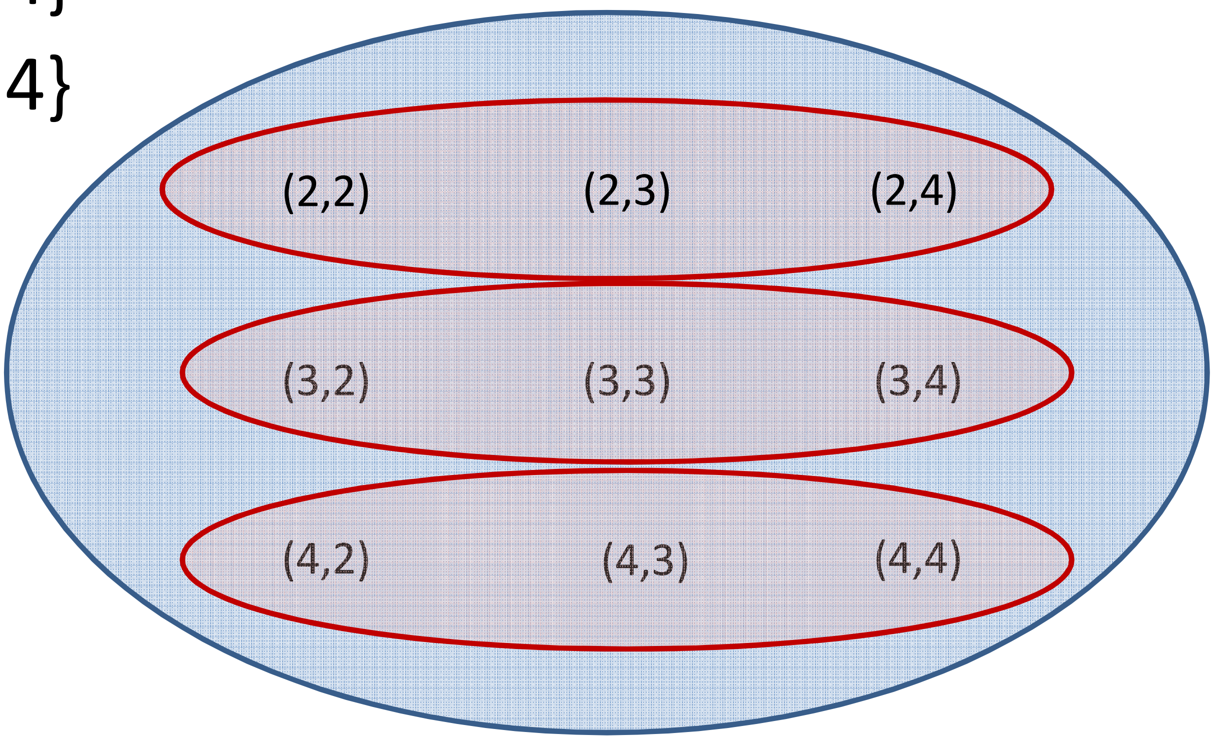
- What is the probability of Whirligig beetles producing 2 offspring in the first litter **and** 3 offspring in the second litter?

Offspring $L_1 = \{2,3,4\}$

Offspring $L_2 = \{2,3,4\}$

$P(L_1 = 2) = 1/3$

$P(L_2 = 3) = 1/3$



Shared events

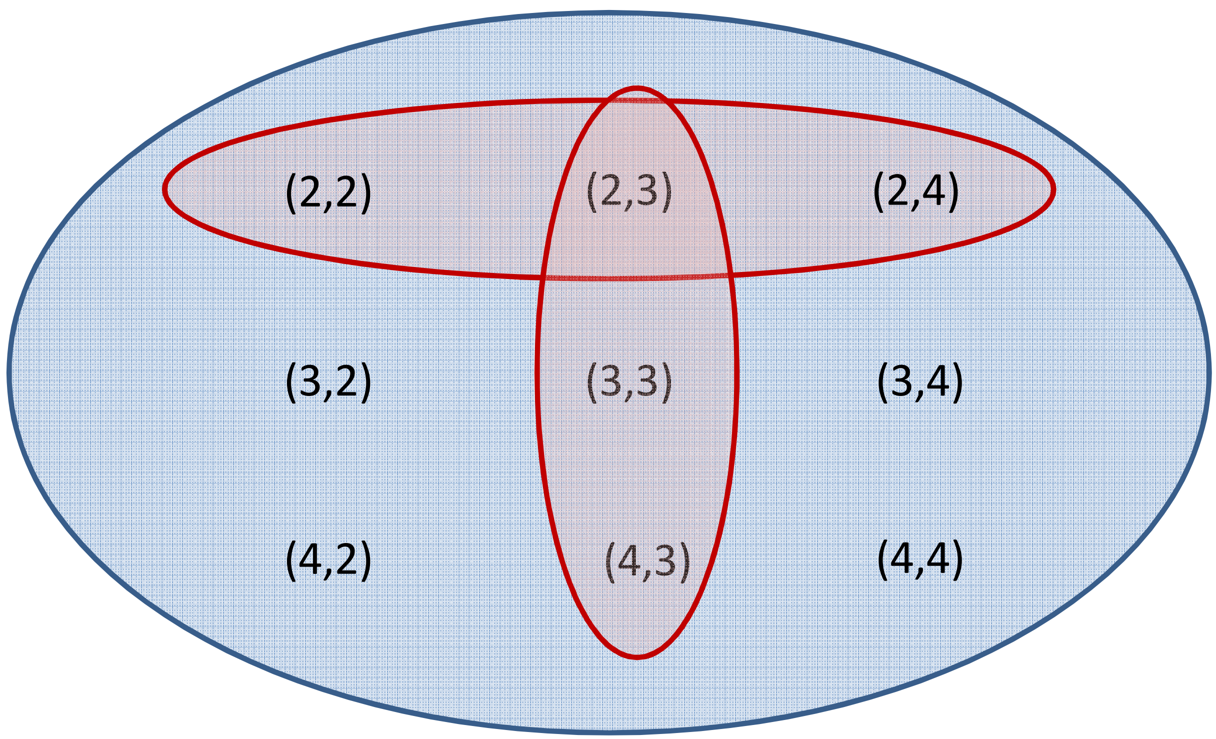
Assume # offspring in second litter is **independent** of first litter

Shared event probability (**intersection** of simple events)
$P(L_1=2 \cap L_2=3) = P(L_1=2) \times P(L_2=3)$

$P(L_1=2 \cap L_2=3) = 1/3 * 1/3 = 1/9$

P((2,3) = 1/9

$P(A \cap B) = P(A) \times P(B)$

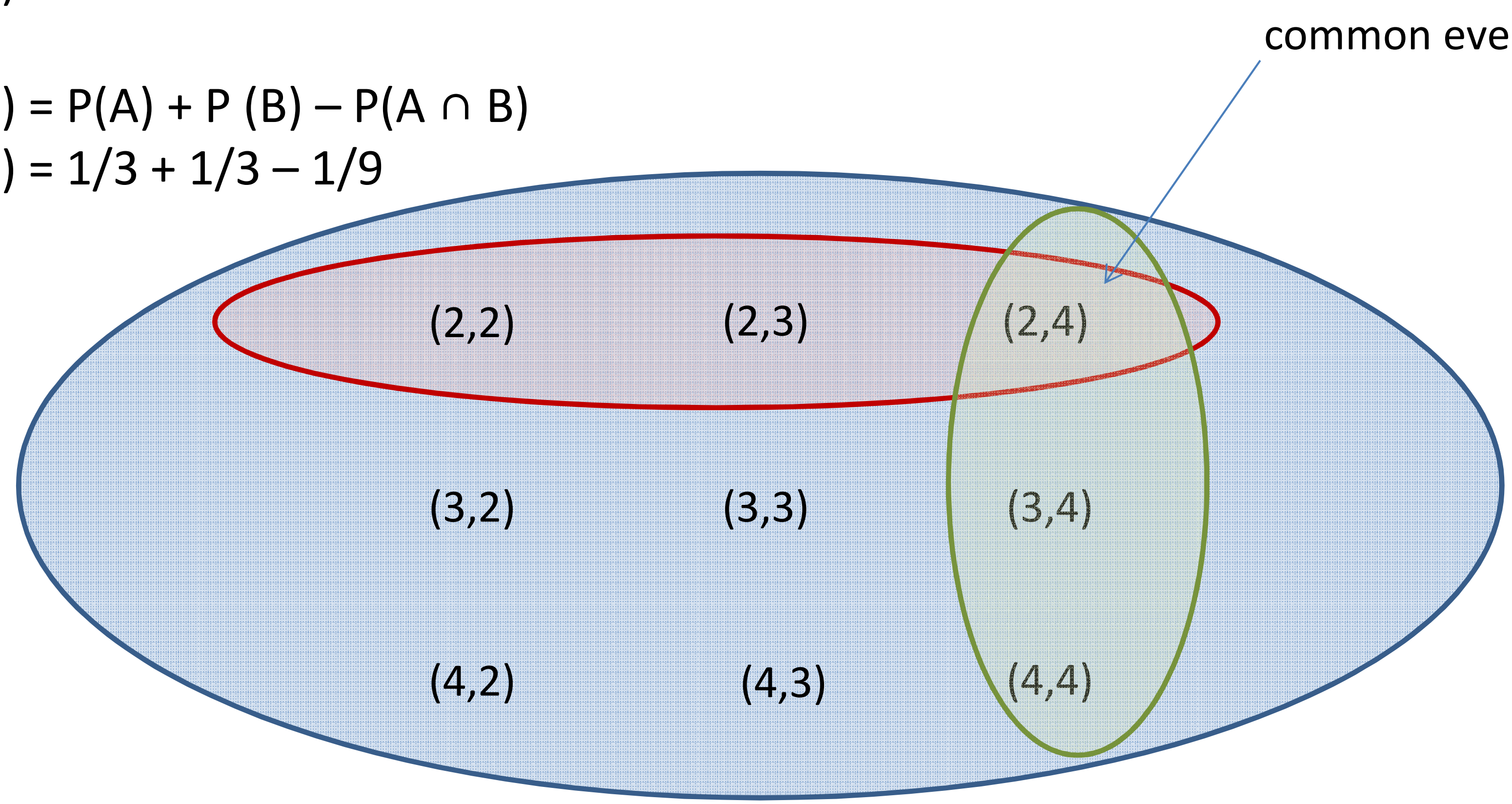


Union of events

Probability of a union of two events
Assume # offspring in second litter is **independent** of first litter

Probability that $L_1=2$ **or** $L_2=4$
$P(L_1=2 \cup L_2=4)$
$P(A \cup B)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$P(A \cup B) = 1/3 + 1/3 - 1/9$



Conditional probability

Events are **NOT** independent

Offspring produced in L2 are dependent on # of offspring in L1

P(A|B) = P(A) x P(B) P(A ∩ B)

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$P(L_2\text{=4 given}L_1 = 2) = \frac{(1/9)}{(1/3)}$