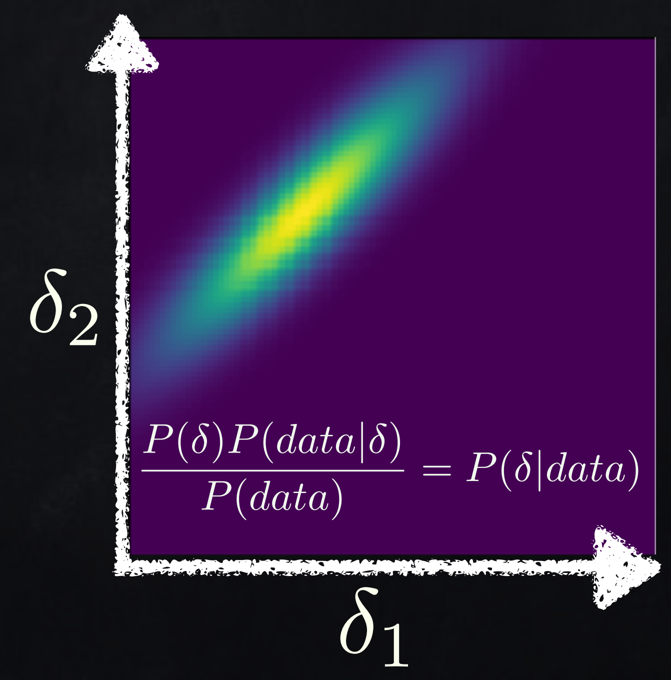
Bayesian inference methods for ecology



Ruben Perez-Carrasco
r.perez-carrasco@imperial.ac.uk

Bayesian inference methods for ecology

Theory

What is Bayesian inference?

Probability 101

Bayes' theorem in biology

Priors and likelihood

Interpretation of posteriors

Markov Chain Monte Carlo

Applications

Population dynamics in microscopic and macroscopic communities

Phylogeny and model selection

ODE generative models: Morphogen patterning of embryonic tissues

Bayesian inference methods for ecology

Theory

What is Bayesian inference?

Probability 101

Bayes' theorem in biology

Priors and likelihood

Interpretation of posteriors

Markov Chain Monte Carlo

At the end of today's class you should...

- Understand what is Bayesian inference and why is important in Biology
- Be confident with the manipulation of probability distributions
- Identify the basic steps in applying Bayesian inference
- Interpret the biological meaning of a posterior distribution

Bayesian inference methods for ecology

Theory

What is Bayesian inference?

Probability 101

Bayes' theorem in biology

Priors and likelihood

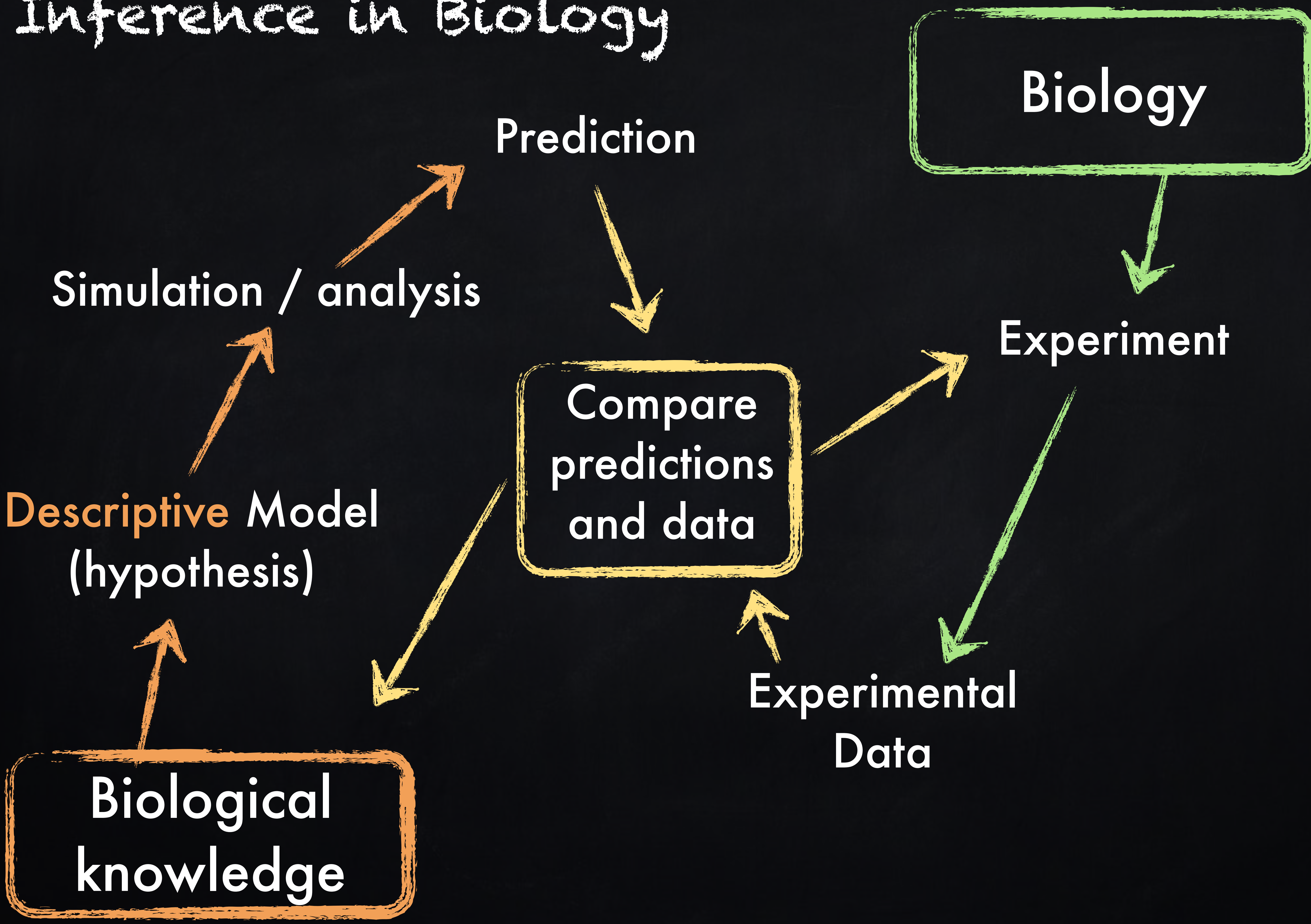
Interpretation of posteriors

Markov Chain Monte Carlo

At the end of today's class you should...

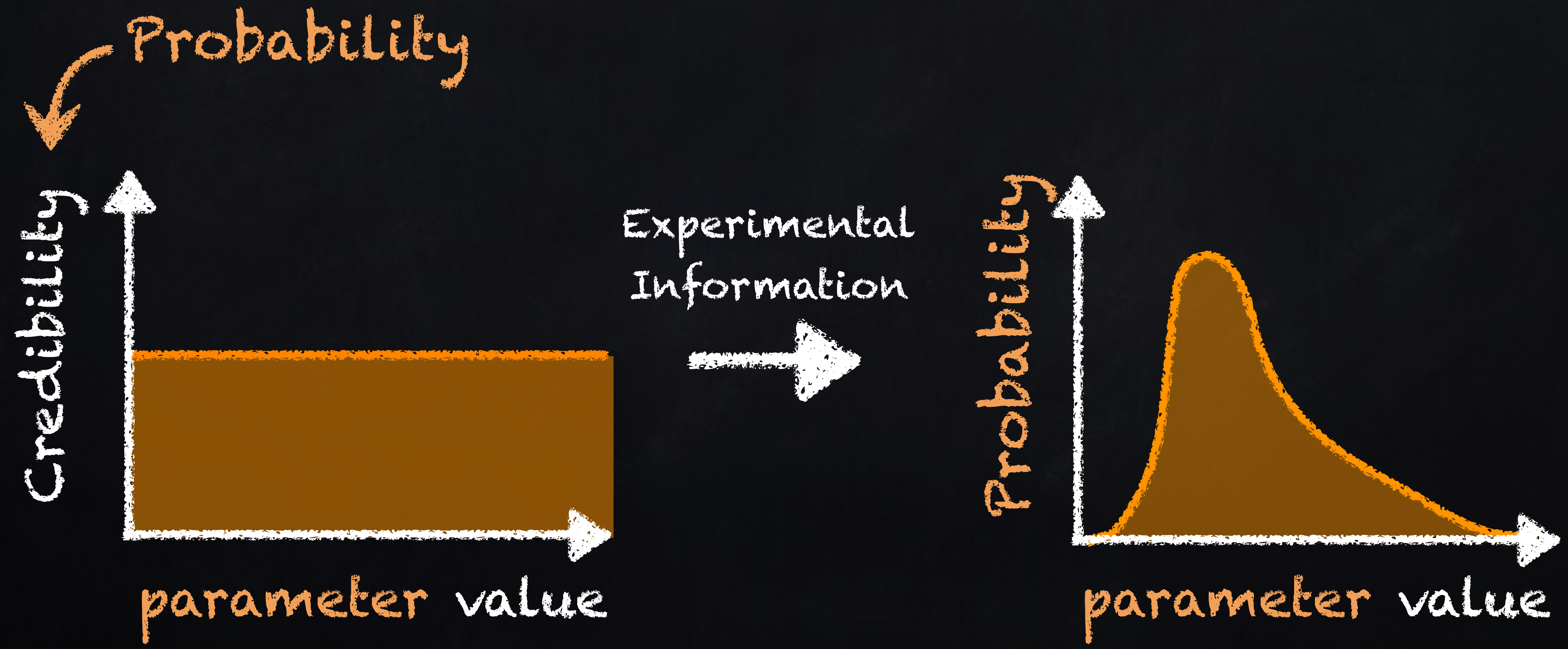
- Understand what is Bayesian inference and why is important in Biology
- Be confident with the manipulation of probability distributions
- Identify the basic steps in applying Bayesian inference
- Interpret the biological meaning of a posterior distribution

Inference in Biology



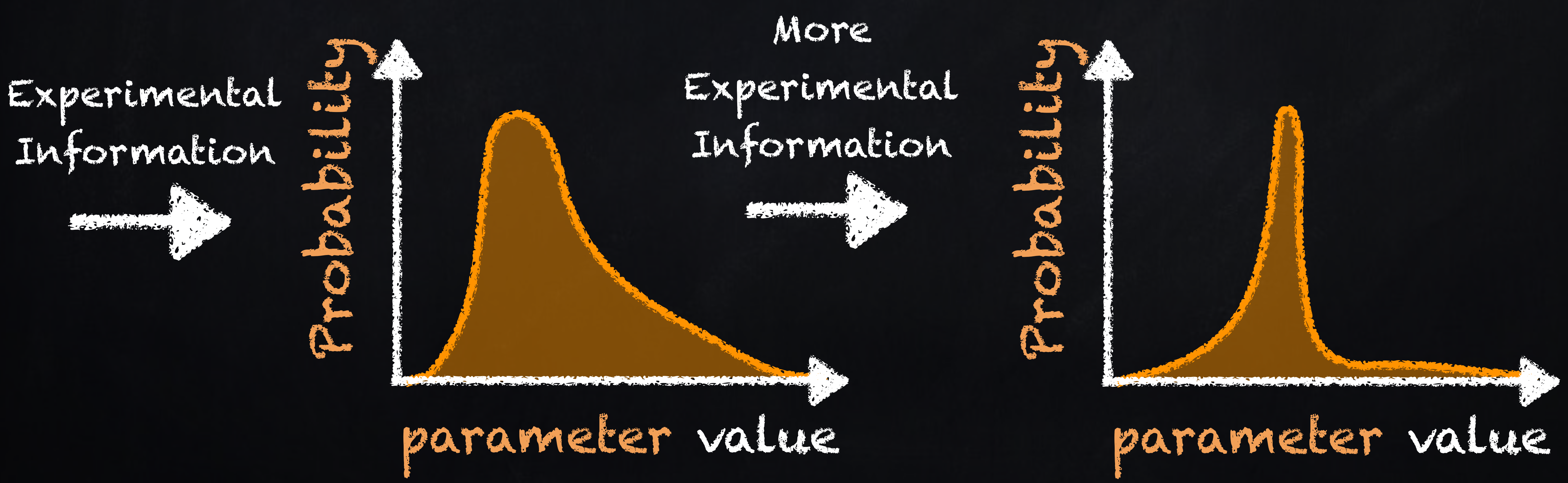
Bayesian inference is …

… reallocation of credibility across possibilities using descriptive mathematical models



Bayesian inference is …

… reallocation of credibility across possibilities using descriptive mathematical models



Bayesian Inference in systems biology

Theory

What is Bayesian inference?

Probability 101

Bayes' theorem in biology

Priors and likelihood

Interpretation of posteriors

Markov Chain Monte Carlo

Applications

Population dynamics in microscopic and macroscopic communities

Phylogeny and model selection

ODE generative models: Morphogen patterning of embryonic tissues

Probability 101

Random variable X: Possible outcome of an observation

Random variables can be discrete or continuous

Discrete Probability $P(X)$

impossible determined

$$0 \leq P(X) \leq 1$$

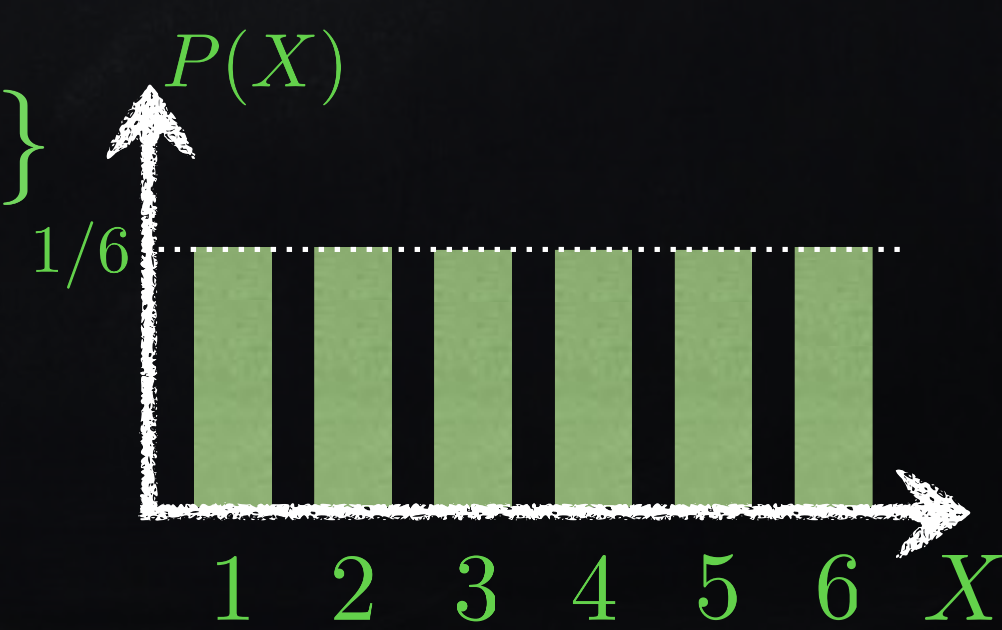
normalisation

$$\sum_{\text{all outcomes}} P(X) = 1$$

$X \in \{1, 2, 3, 4, 5, 6\}$

X is discrete

$$P(X = 3) = \frac{1}{6}$$



Probability 101

Random variable X: Possible outcome of an observation

Random variables can be discrete or continuous

Discrete Probability $P(X)$

impossible $\quad$ determined

$$0 \leq P(X) \leq 1$$

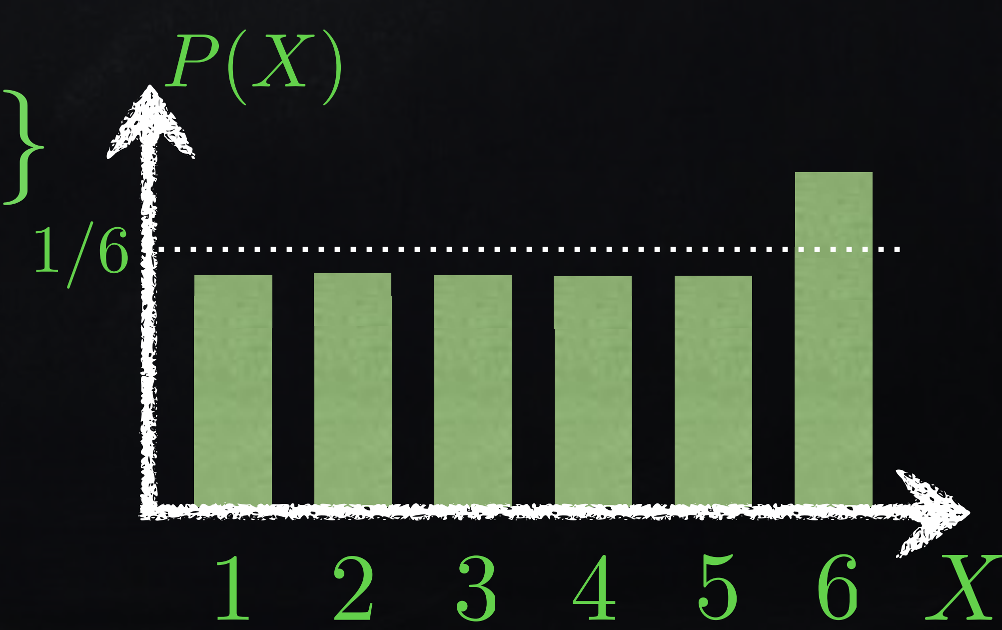
normalisation

$$\sum_{\text{all outcomes}} P(X) = 1$$

$$X \in \{1, 2, 3, 4, 5, 6\}$$

X is discrete

$$P(X = 3) = \frac{1}{6}$$



Probability 101

Random variable X: Possible outcome of an observation

Random variables can be discrete or continuous

Discrete Probability $P(X)$

impossible

determined

$$0 \leq P(X) \leq 1$$

normalisation

$$\sum_{\text{all outcomes}} P(X) = 1$$

Common discrete distributions

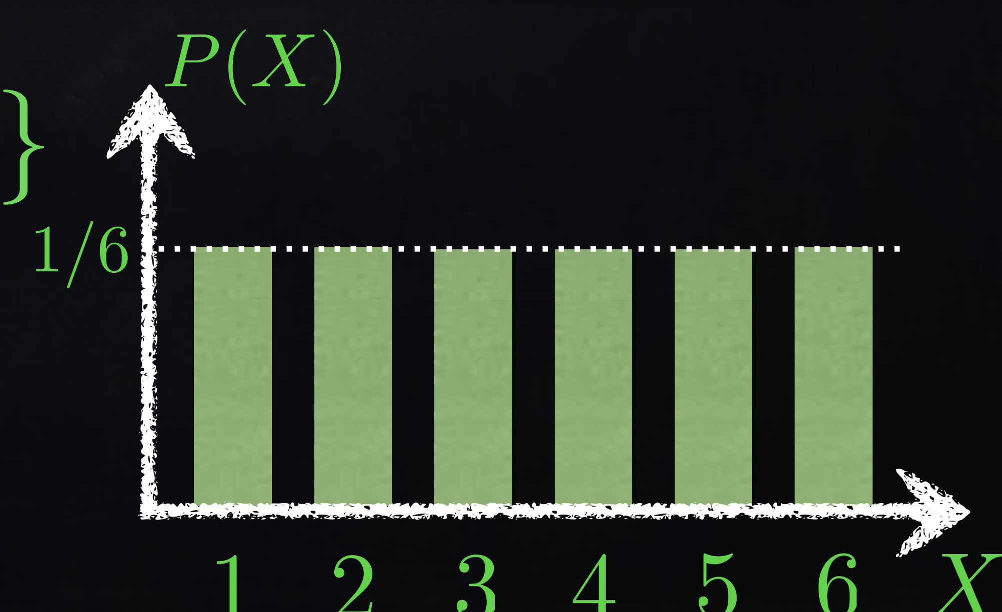
- Uniform $P(X;N) = \frac{1}{N}$
- Poisson $P(X;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$
- Binomial

$$P(X;p,N) = \binom{N}{X} p^X (1-p)^{N-X}$$

$$X \in \{1,2,3,4,5,6\}$$

X is discrete

$$P(X=3) = \frac{1}{6}$$



Probability 101

Random variable X : Possible outc

Random variables can b

Discrete Probability $P(X$ ... ns

impossible determi

$$0 \leq P(X) \leq 1$$

normalisation

$$\sum_{\text{all outcomes}} P(X) = 1$$

$X \in \{1, 2,$

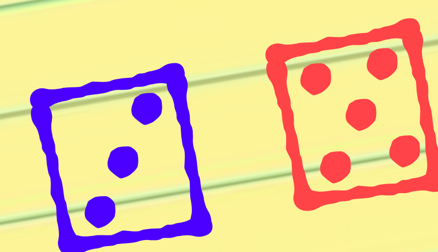
X is discre

$P(X = 3) =$

Question:

You have two regular 6-faced dice, one blue and one red.

What is the probability of throwing them and having the following result?



What is the probability of throwing both and getting a "3" in only one of them?
And at least one of them?

Probability 101

Random variable X: Possible outcome of an observation

Random variables can be discrete or continuous

Continuous probability $X \in \mathbb{R}$

$$\text{Probabibility}(A < X < B) = \int_A^B P(X)\mathrm{d}X$$

Probability density function (PDF)

$$\int_{\min X}^{\max X} P(X)\mathrm{d}X = 1$$

normalisation

Common distributions

$$P(X, N) = constant$$

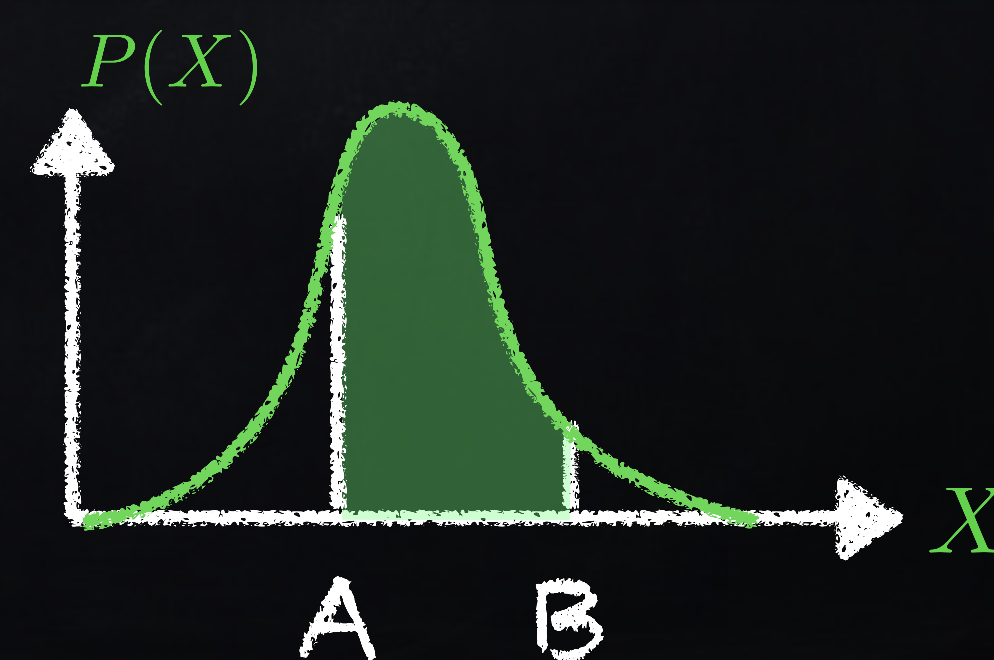
$$P(X; k) = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}$$

- Exponential

$$P(X; \lambda) = \lambda \mathrm{e}^{-\lambda X}$$

$X \in [0, 80\ km/h]$

X is continuous



Probability 101

Random variable X: Possible outcome of an observation

Random variables can be discrete or conti

Continuous probab ... utions

Probabibility$(A < X$

ant

$-\frac{1}{2}(\frac{X-\mu}{\sigma})^2$

$$\int_{\min X}^{\max X} P(X)\mathrm{d}X = 1$$

normalisation

$X \in [$

X is c

X

Question:

Find the value of A so the the following distributions are PDFs

$$P(x) = A, \ \in [-1/2, 1/2]$$

$$P(x) = Ae^{-x}, \ \ x \in [0, \infty]$$

Probability 101

2 random variables: X, Y

Joint Probability	$P(X, Y)$	Probability of an event from X AND an event from Y
Conditional Probability	$P(X\|Y)$	Probability of an event from X GIVEN an event from Y
Marginal Probability	$P(X)$	Probability of an event from X INDEPENDENTLY from Y

$$\boxed{P(X,Y) = P(X)P(Y|X)}$$

X = month $\quad$ Y = day number

P(November, 7) = 1/366

P(November | 7) = 1/12

P(7) = 12/365

Probability 101

2 random variabl

Joint Probability

Conditional Probability

Marginal Probability

rom X

om X

m X

$P(X, Y) =$

X = mo

P(Nover

P(Noven

$P(7) =$ 1

Question:

- Calculate the following probabilities:

$P(31, \text{December})$

$P(31)$

$P(\text{Dec})$

$P(31 | \text{December})$

$P(\text{December} | 31)$

and test that:

$P(31, \text{Dec}) = P(31 | \text{Dec}) P(\text{Dec})$

$P(31, \text{Dec}) = P(\text{Dec} | 31) P(31)$

Probability 101

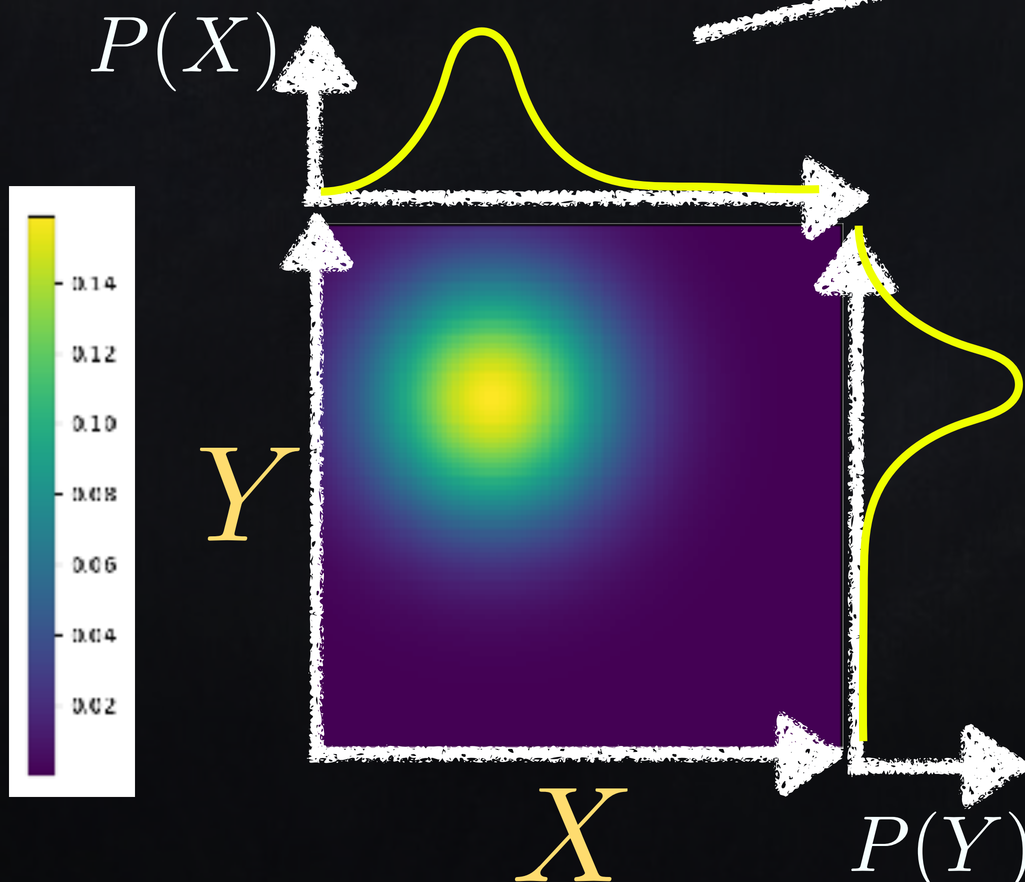
2 continuous random variables: X, Y

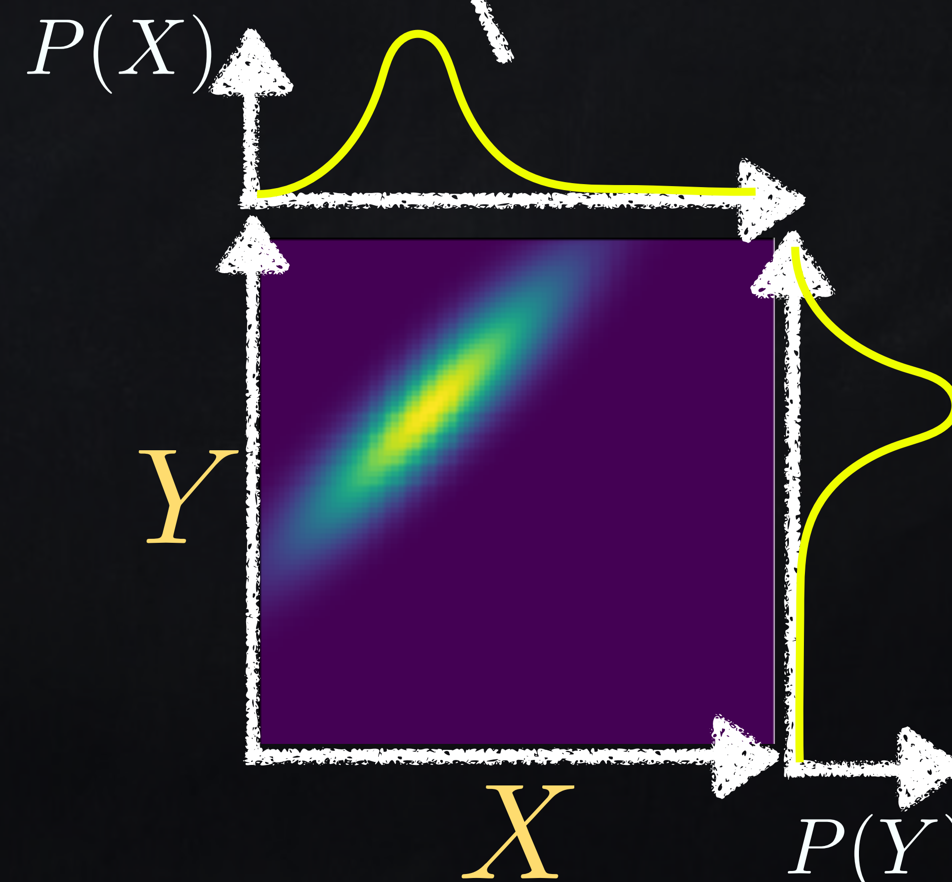
$$P(X,Y) = P(X)P(Y|X)$$

$$\int_{minY}^{maxY} P(X,Y)\mathrm{d}Y$$

Same marginals but different joint distributions!

$P(X,Y)$





Independent variables

$$P(Y|X) = P(Y)$$
$$P(X,Y) = P(X)P(Y)$$

Dependent variables

$$P(Y|X) \neq P(Y)$$
$$P(X,Y) \neq P(X)P(Y)$$

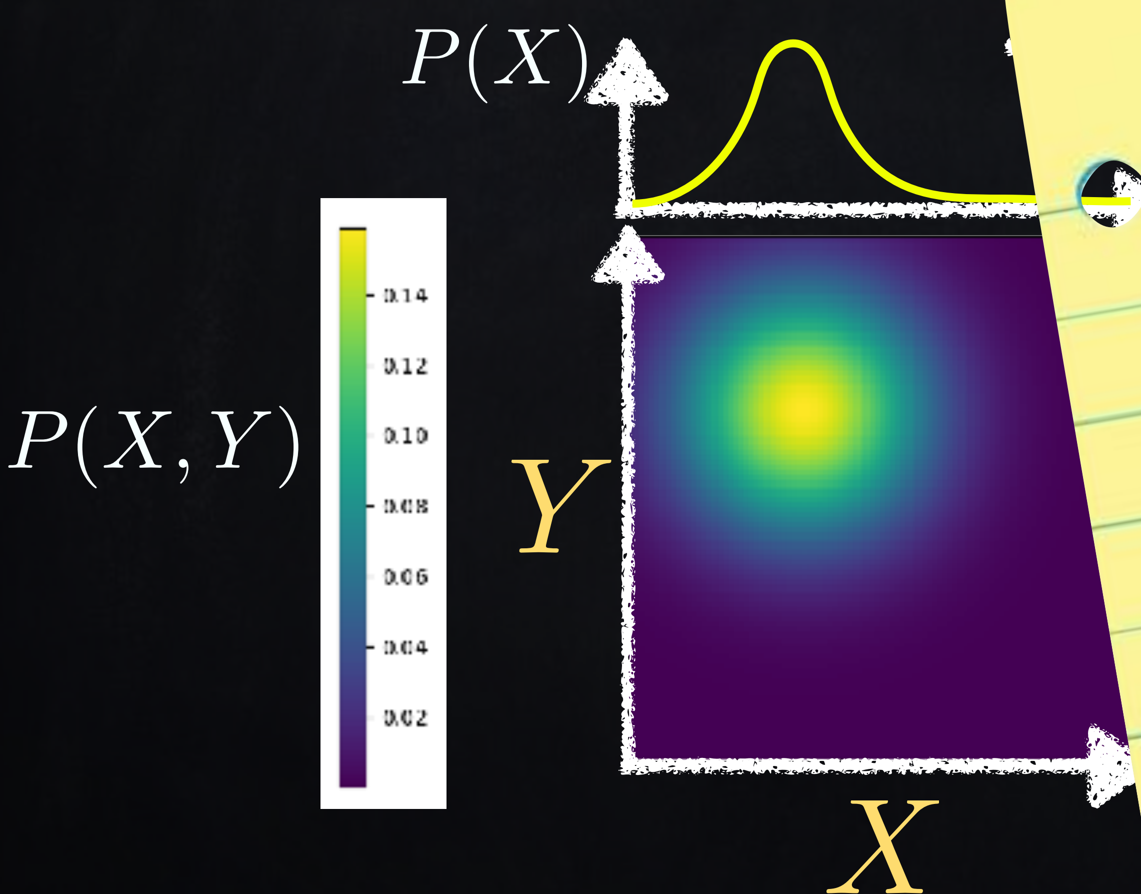
Probability 101

2 continuous random variables: X, Y

$$P(X,Y) = P(X)P(Y|X)$$

$$\int_{minY}^{maxY} P(X,Y)\mathrm{d}Y$$

Same marginals but
different joi !



Independent variables

$$P(Y|X) = P(Y)$$

$$P(X,Y) = P(X)P(Y)$$

Question:

- Can we have a situation where X is a discrete random variable and Y a continuous one?

- What is the normalisation condition for P(X,Y) where X and Y are continuous random variables?

Probability 101

Time t is usually not a random variable

Probability in time $P(X,t)$

t can be discrete (e.g. generations) or continuous (e.g. duration)

Steady state probability $P(X, t \to \infty) = \tilde{P}(X)$

t signifies sequential ordering i.e. causality in our model

Generative model:

$$P(X,t_N) = F(P(X,t_{N-1}), P(X,t_{N-2}), \dots, P(X,t_0))$$

Special case (Markov process)

$$P(X,t_N) = F(P(X,t_{N-1}))$$

e.g. Population dynamics,
Reaction kinetics
Epidemic models

Question:

In the simplest model of DNA substitution (JC69), each nucleotide can mutate to any other nucleotide with the same probability per unit time r.

Knowing that at a given time $t=0$ the base of site is A: $P(A,t=0)=1$. Derive the general expression for the probability $P(G,t)$ of the site having a base G at time t

1) Express the probability that the site has a base G at time $t+\Delta t$ as $P(G|t+\Delta t)=P(G|t)+F(t)$, where $F(t)$ represents the change in probability over the small time interval Δt.

2) Transform the above expression into a differential equation that describes the rate of change of $P(G,t)$ over time.

3) Solve the differential equation to find $P(G,t)$ expression for $P(G,t)$.

Bayesian Inference in systems biology

Theory

What is Bayesian inference?

Probability 101

Bayes' theorem in biology

Priors and likelihood

Interpretation of posteriors

Markov Chain Monte Carlo

Applications

Population dynamics in microscopic and macroscopic communities

Phylogeny and model selection

ODE generative models: Morphogen patterning of embryonic tissues

Bayes' Theorem

$$P(X,Y) = P(X)P(Y|X)$$
$$P(X,Y) = P(Y)P(X|Y)$$

$$\frac{\cancel{P(X,Y)}}{\cancel{P(X,Y)}} = \frac{P(X)P(Y|X)}{P(Y)P(X|Y)}$$

$$\boxed{P(X|Y) = \frac{P(X)P(Y|X)}{P(Y)}}$$

Bayesian Inference

$$P(X|Y)=\frac{P(X)P(Y|X)}{P(Y)}$$

Likelihood

Prior probability

$$\frac{P(\delta)P(data|\delta)}{P(data)} = P(\delta|data)$$

Posterior probability

Y : Experimental data

X : Parameter value (δ)

Bayesian Inference

$$P(X|Y)=\frac{P(X)P(Y|X)}{P(Y)}$$

Likelihood

Prior probability

Posterior probability

$$\frac{P(\delta)P(data|\delta)}{P(data)} = P(\delta|data)$$

What about $P(data)$? ✓

How do I choose the prior? ✓

What am I supposed to do with the posterior? ✓

Okay, but how do apply this with my laptop? ✓

Can you give examples of the likelihood? ✓