A very small intro to MCMC

Rosana Zenil-Ferguson

University of Hawai'i

December 2019

Markov Chain Model

Markov Chain Model

A stochastic process in time $X(t)$ in which the probability of what happens next depends only on the current state, and it is not affected by additional knowledge of the past.

Markov Chain Model

A stochastic process in time $X(t)$ in which the probability of what happens next depends only on the current state, and it is not affected by additional knowledge of the past.

Discrete time Markov chains (DTMC)

$$P(X_t|X_{t-1}, X_{t-2}, \ldots, X_0) = P(X_t|X_{t-1})$$

Continuous time Markov chains (CTMC)

$$P(X_t|X_s, X_r) = P(X_t|X_s) \quad \text{when } r < s$$

Markov chain in the context of MCMC

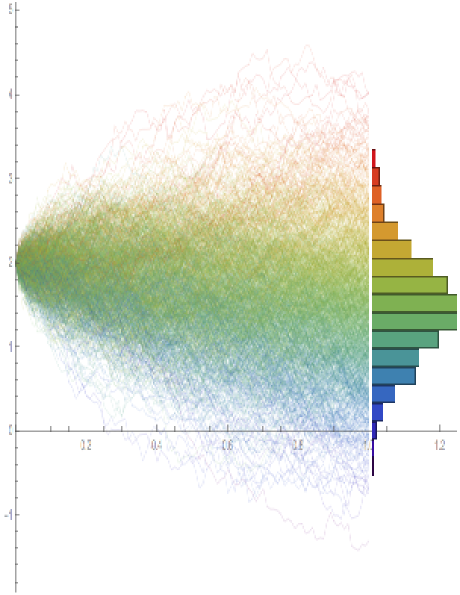
- When modeling in phylogenetics we used CTMC because of the Markovian property.

Markov chain in the context of MCMC

- When modeling in phylogenetics we used CTMC because of the Markovian property.
- When using a Markov Chain for inference(MCMC) we use DTMC and we are also interested in the **ergodic property**

An example of ergodic Markov Chain: Brownian Motion

Stationary distribution $N(\mu, \sigma^2 t)$



What we want

To know the Posterior distribution $P(\theta|D)$

What we want

To know the Posterior distribution $P(\theta|D)$

Goal: To build a Markov chain that in the long run has $P(\theta|D)$ as the stationary distribution.

Metropolis-Hastings algorithm. Intuition

- If we some a set of parameters θ with probability $P(\theta|D)$

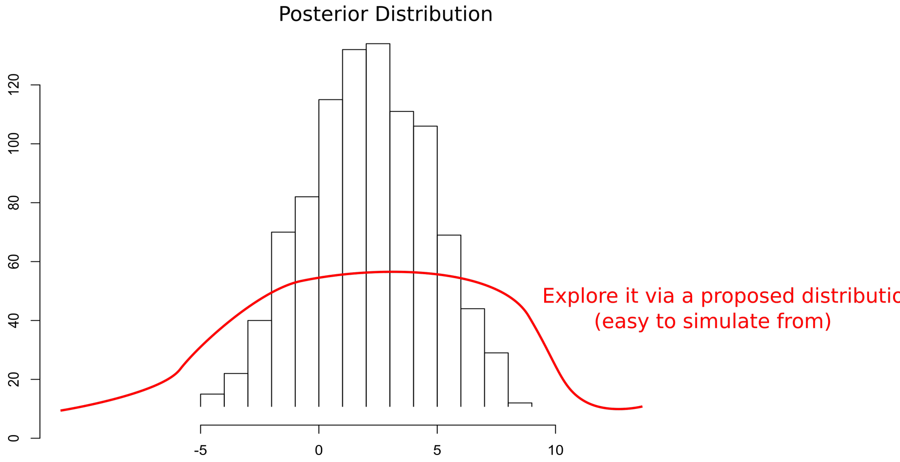
Metropolis-Hastings algorithm. Intuition

- If we some a set of parameters θ with probability $P(\theta|D)$
- We want to estimate new parameters θ' such that the posterior odds are high $\frac{P(\theta'|D)}{P(\theta|D)}$

Metropolis-Hastings algorithm. Intuition

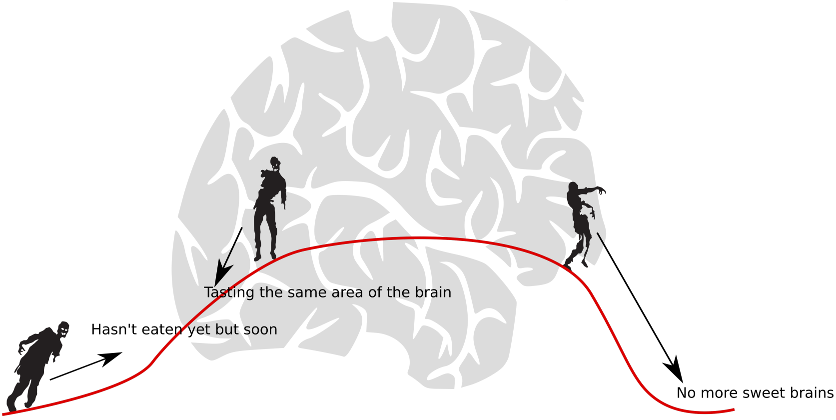
- If we some a set of parameters θ with probability $P(\theta|D)$
- We want to estimate new parameters θ' such that the posterior odds are high $\frac{P(\theta'|D)}{P(\theta|D)}$
- So how do we propose those new parameters θ'?

Answer: A proposal distribution that is easy to simulate from



Proposal moves explained via zombies

Zombie Goal: To eat all the brain as fast as possible



Based on Paul Lewis' Bayesian Statistics lectures, 2017.

Metropolis-Hastings algorithm

1. Select an initial value of θ_0

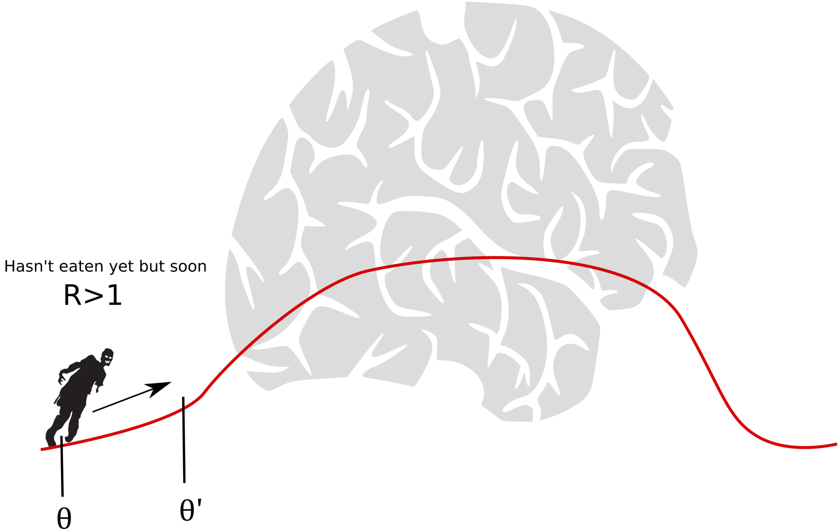
Metropolis-Hastings algorithm

1. Select an initial value of θ_0
2. A new value of the parameters θ' is a draw from proposal $q(\theta'|\theta)$

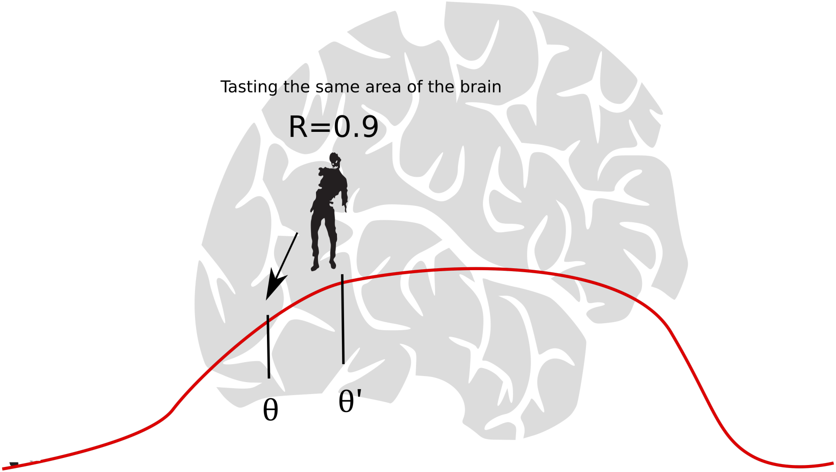
Metropolis-Hastings algorithm

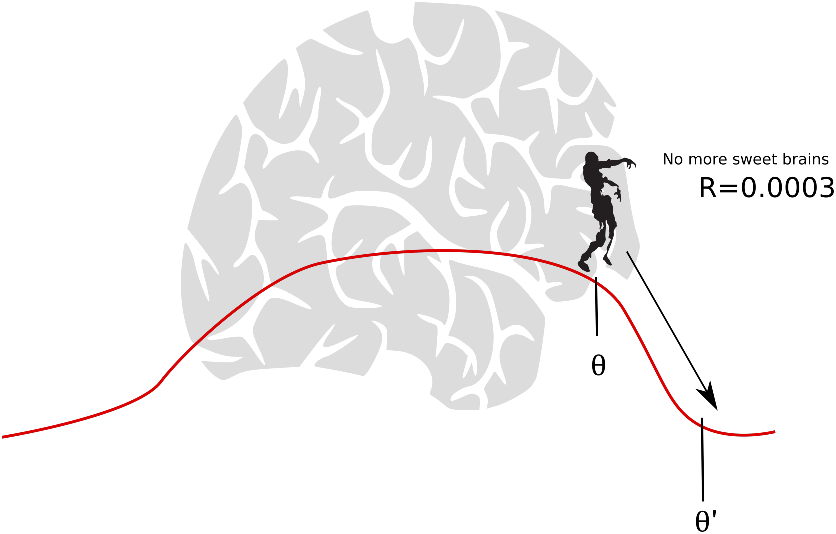
1. Select an initial value of θ_0
2. A new value of the parameters θ' is a draw from proposal $q(\theta'|\theta)$
3. Is that new value good? It depends on the **Acceptance ratio** R: The posterior odds and the "direction"

$$R = \frac{P(\theta'|D)}{P(\theta|D)} \frac{q(\theta|\theta')}{q(\theta'|\theta)}$$



Hasn't eaten yet but soon
R>1
θ
θ'


Tasting the same area of the brain
R=0.9
θ
θ'


No more sweet brains
R=0.0003
θ
θ'

Metropolis-Hastings algorithm

1. Select an initial value of θ_0
2. A new value of the parameters θ' is a draw from proposal $q(\theta'|\theta)$
3. Draw a uniform value u between (0,1). If $u < min(1, R)$ then move to θ'

Additional considerations of the acceptance ratio

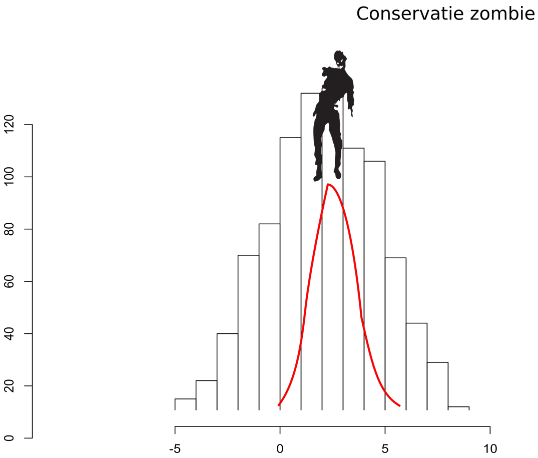
Acceptance ratio R: The posterior odds and the "direction" is usually written in likelihood-prior form

$$R = \frac{P(\theta'|D)}{P(\theta|D)} \frac{q(\theta|\theta')}{q(\theta'|\theta)}$$

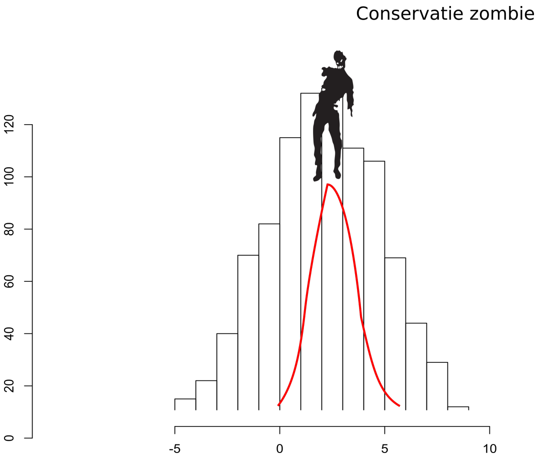
$$R = \frac{P(D|\theta')P(\theta')}{P(D|\theta)P(\theta)} \frac{q(\theta|\theta')}{q(\theta'|\theta)}$$

And it simplifies if the proposal is symmetric $q(\theta'|\theta) = q(\theta|\theta')$

Let's think further about proposals

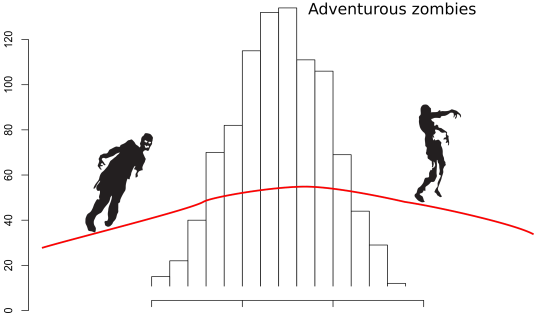


Let's think further about proposals

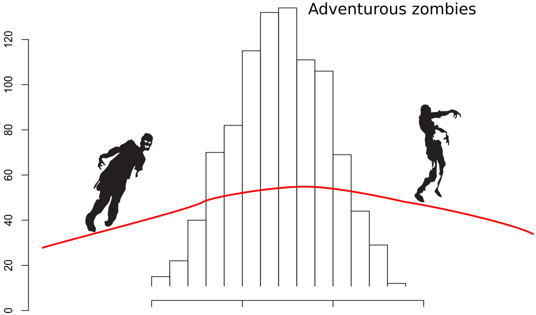


- Pro: Zombie will be constantly eating brain
- Con: Zombie will take forever to eat the whole brain (or will never finish eating)

Let's think further about proposals



Let’s think further about proposals



- Pro: Zombie will finish eating the brain
- Con: Zombie will go through long periods of brain shortage