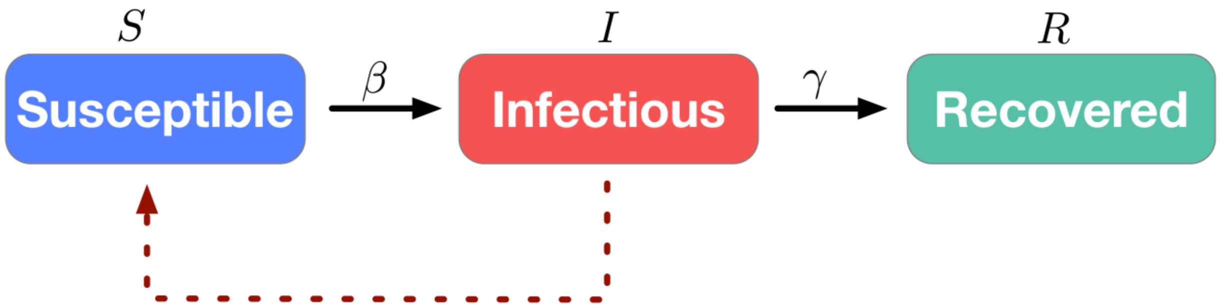
Introduction to MCMC for Infectious Diseases

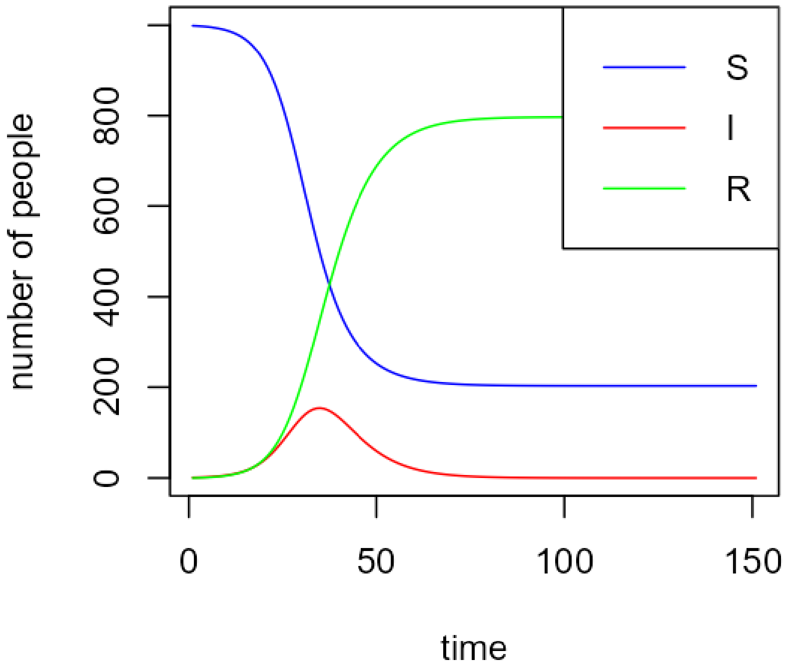
CCB Skills Seminar – 10/10/2024

Philip Lee

Long history of modeling infectious disease transmission

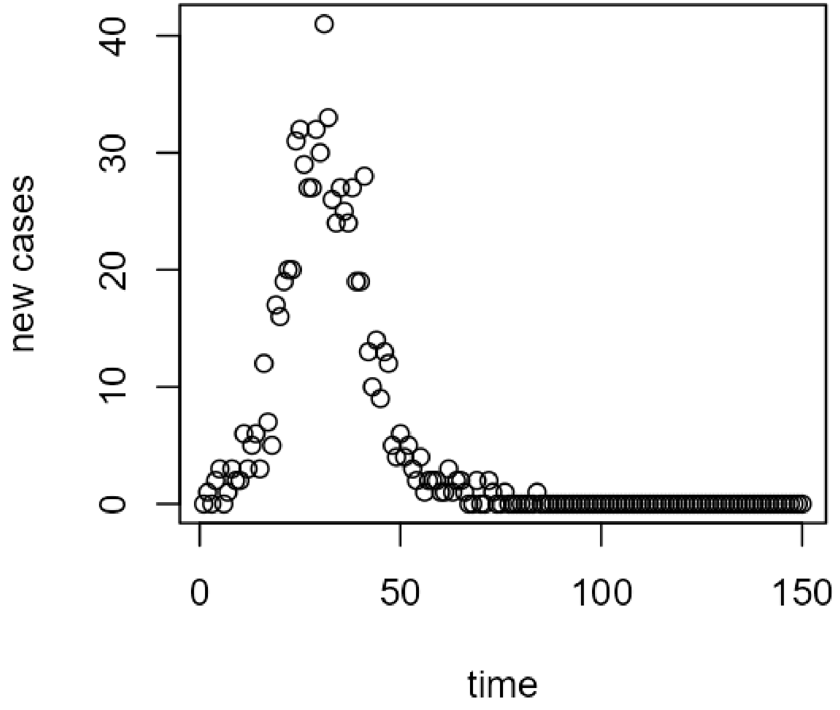


$$\frac{dS}{dt} = -\beta SI/N$$
$$\frac{dI}{dt} = \beta SI/N - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$



SIR schematic source: https://covid19.uclaml.org/model.html

Given data, what can we say about our model parameters?



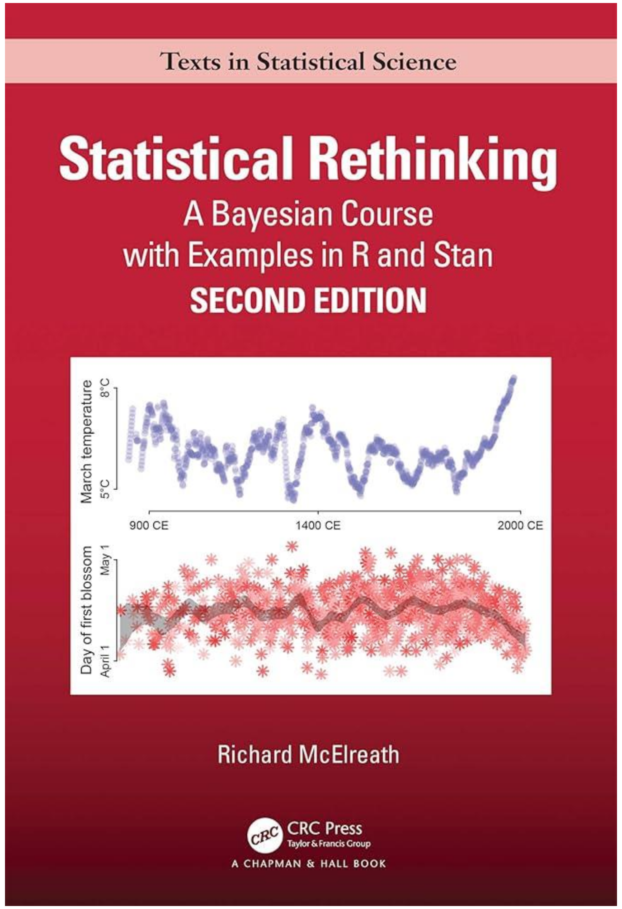
?

$$\frac{dS}{dt} = -\beta SI/N$$

$$\frac{dI}{dt} = \beta SI/N - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

Resource for Bayesian statistics



Bayesian inference

- **Observed data:** e.g., number of observed infections (D)

- **Probability model:** how is the observed data generated?
  - Model is parameterized by $\theta \in \Theta$

- **Fitting the model:** we want to find θ that have high probability given our observed data
  - We want $P(\theta|D)$ to be large
  - i.e., we want to find "highly plausible" θ given our data

Bayesian updating

- **Bayes' rule** tells us how to compute these "plausibilities":

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

- $P(\theta)$ is the **prior probability**
- $P(D|\theta)$ is the **likelihood**
- $P(D)$ is a normalizing constant
- $P(\theta|D)$ is the **posterior probability**

Globe tossing example

- Toss globe 9 times: observe $W = 6$ water and $L = 3$ land
- What's the proportion p of Earth that is covered in water?

- Observed data $D = (W, L)$
- Binomial model with parameter p
  - $W \sim \mathrm{Binom}(p)$
- **Question:** which p are most plausible given our data?

Globe tossing Bayesian updating

- Apply Bayes' rule:

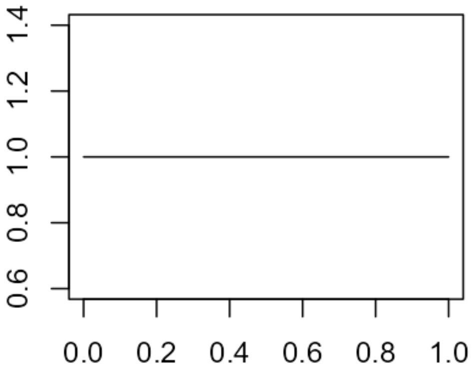
$$P(p|W) = \frac{P(W|p)P(p)}{P(W)} \qquad P(p|W) \propto P(W|p)P(p)$$

- **Binomial likelihood:** $P(W|p) = \binom{W+L}{W} p^W (1-p)^L$

- **Uniform prior:** $P(p) = 1$

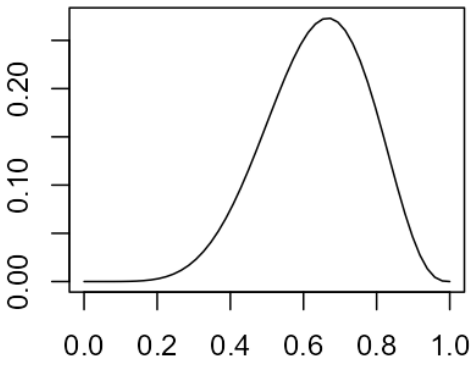
Globe tossing Bayesian updating

Prior



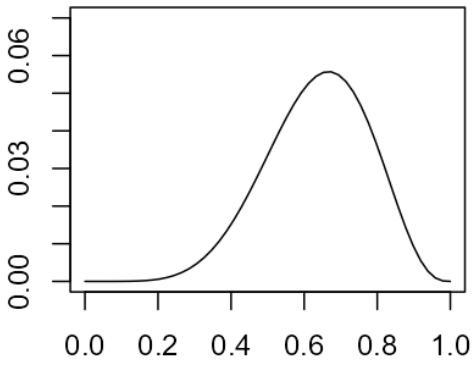
$\times$

Likelihood

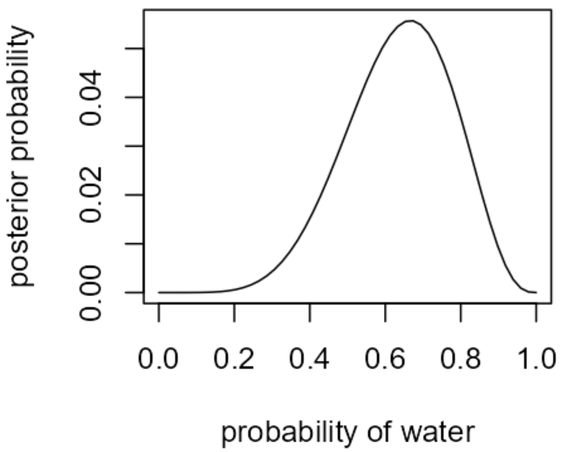


$\propto$

Posterior



Analyzing the Bayesian estimate



- The **entire** posterior distribution is the Bayesian estimate!

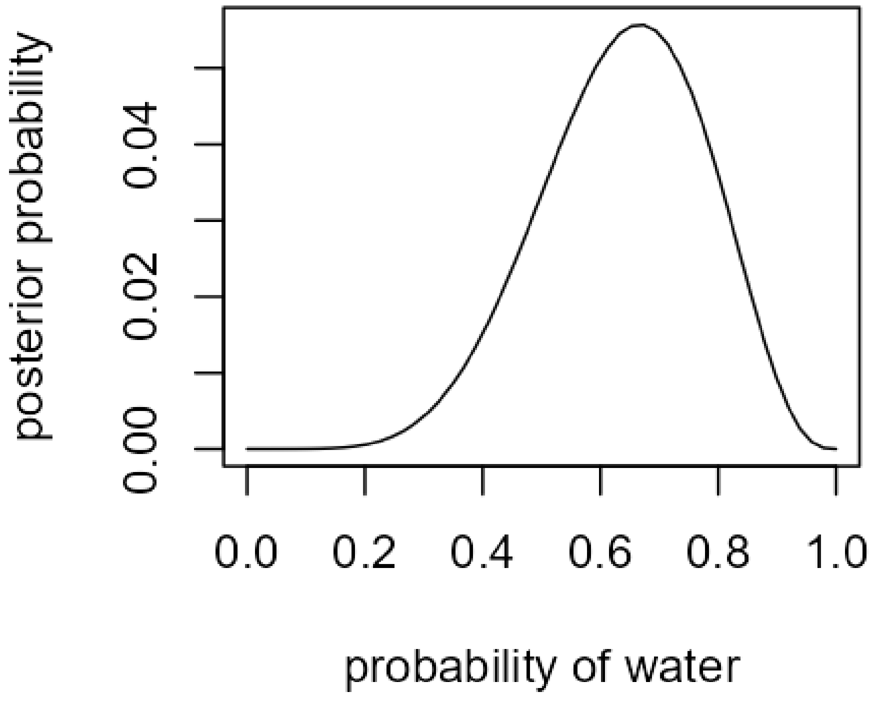
- Extract **point estimates**
  - Find the parameter value with highest probability (MAP estimate)
  - Find the median parameter value (50% quantile)

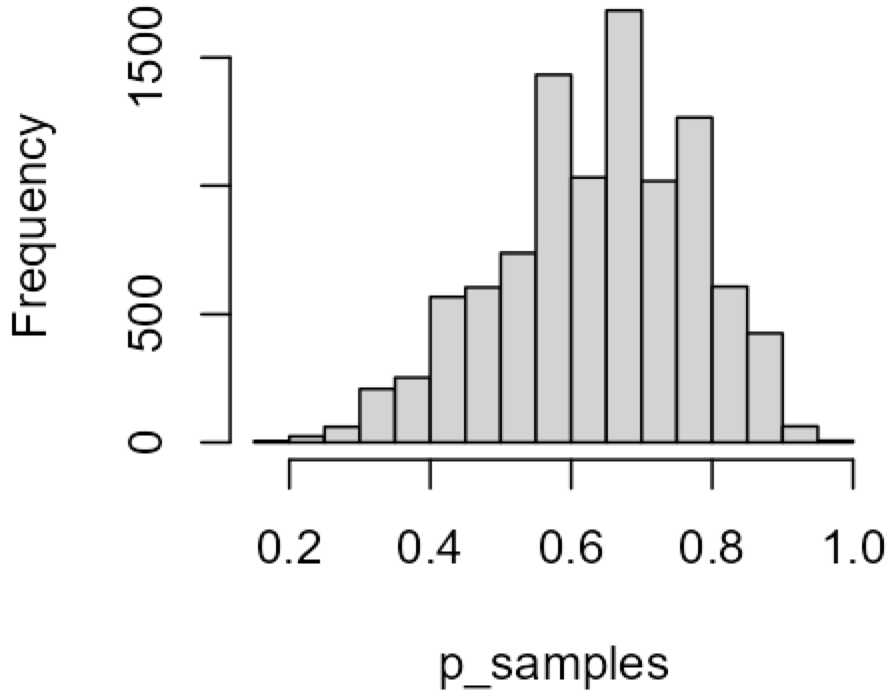
- Extract **credible intervals**
  - e.g., find the 2.5% and 97.5% quantiles of the parameter values

- Make **posterior predictions**
  - Simulate new data taking into account all of the uncertainty embodied in the posterior distribution

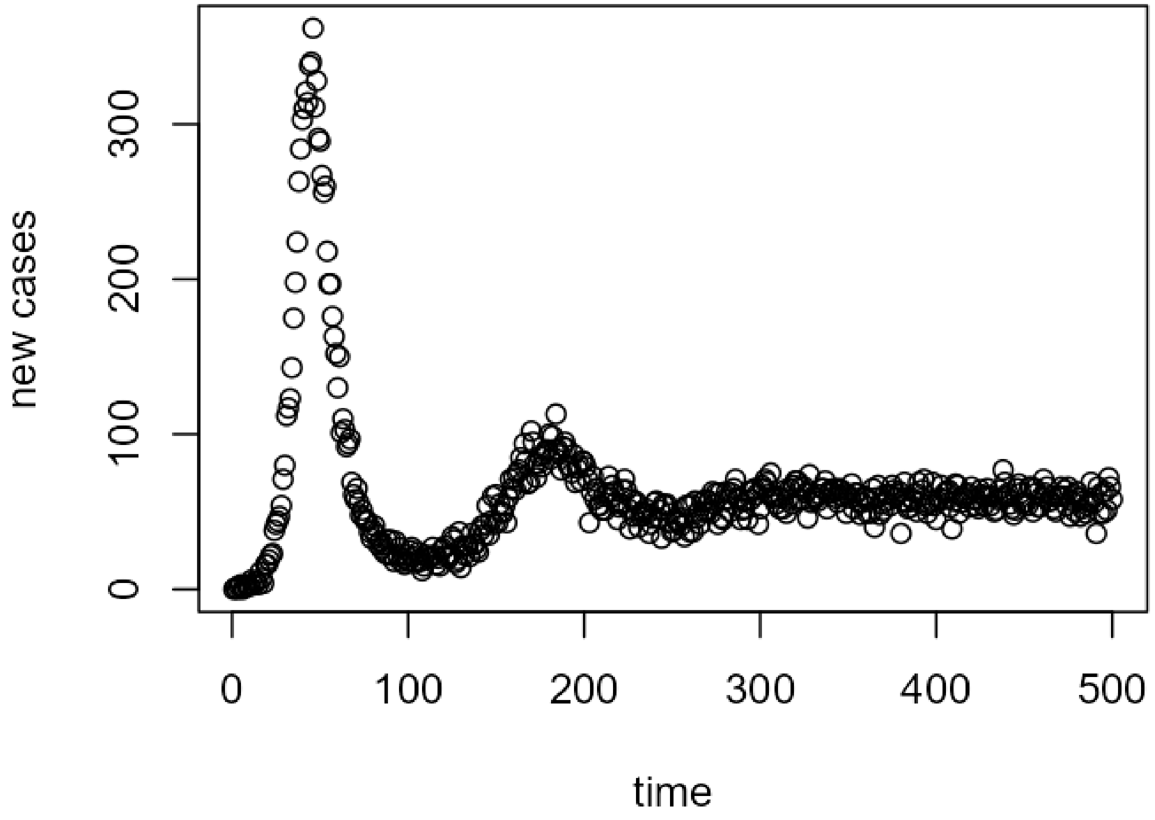
Sampling from the Bayesian estimate

- Usually easier to work with samples from the posterior

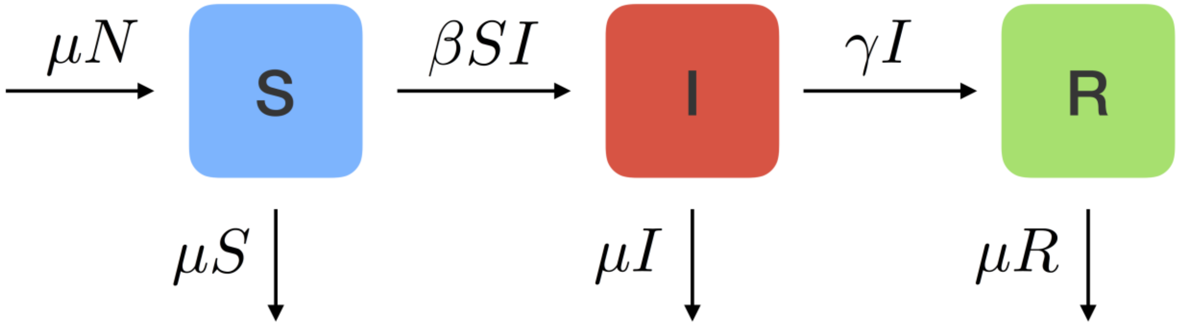




What can we learn from this data?



Need to add demographic processes to obtain periodic case trajectories



$$\frac{dS}{dt} = \mu N - \beta SI/N - \mu S$$
$$\frac{dI}{dt} = \beta SI/N - \gamma I - \mu I$$
$$\frac{dR}{dt} = \gamma I - \mu R$$

SIR schematic source: https://sineadmorris.github.io/post/the-sir-model/

Key epidemiological parameter: R_0

- Average number of secondary cases that an infectious individual would generate in a completely susceptible population
  - $R_0 < 1$: no outbreak
  - $R_0 > 1$: yes outbreak

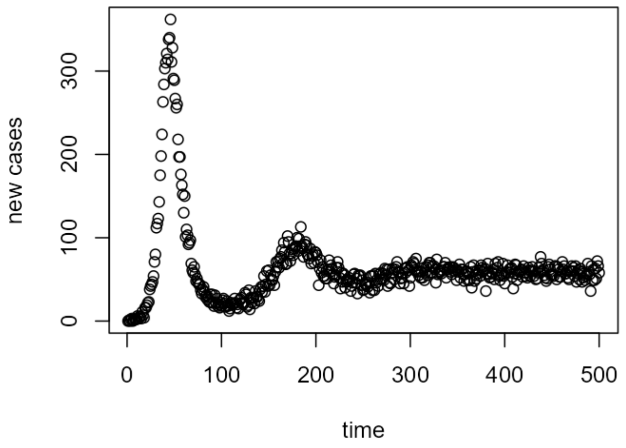
$$\frac{dS}{dt} = \mu N - \beta SI/N - \mu S$$
$$\frac{dI}{dt} = \beta SI/N - \gamma I - \mu I$$
$$\frac{dR}{dt} = \gamma I - \mu R$$

→

$$R_0 = \frac{\beta}{\gamma + \mu}$$

Perform inference on R_0

Observed data D



Probability model
(assume cases Poisson distributed)

$$\frac{dS}{dt} = \mu N - \beta SI/N - \mu S$$
$$\frac{dI}{dt} = \beta SI/N - \gamma I - \mu I$$
$$\frac{dR}{dt} = \gamma I - \mu R$$

Likelihood: $P(D|R_0)$

Prior: $P(R_0)$

Grid approximation can fail

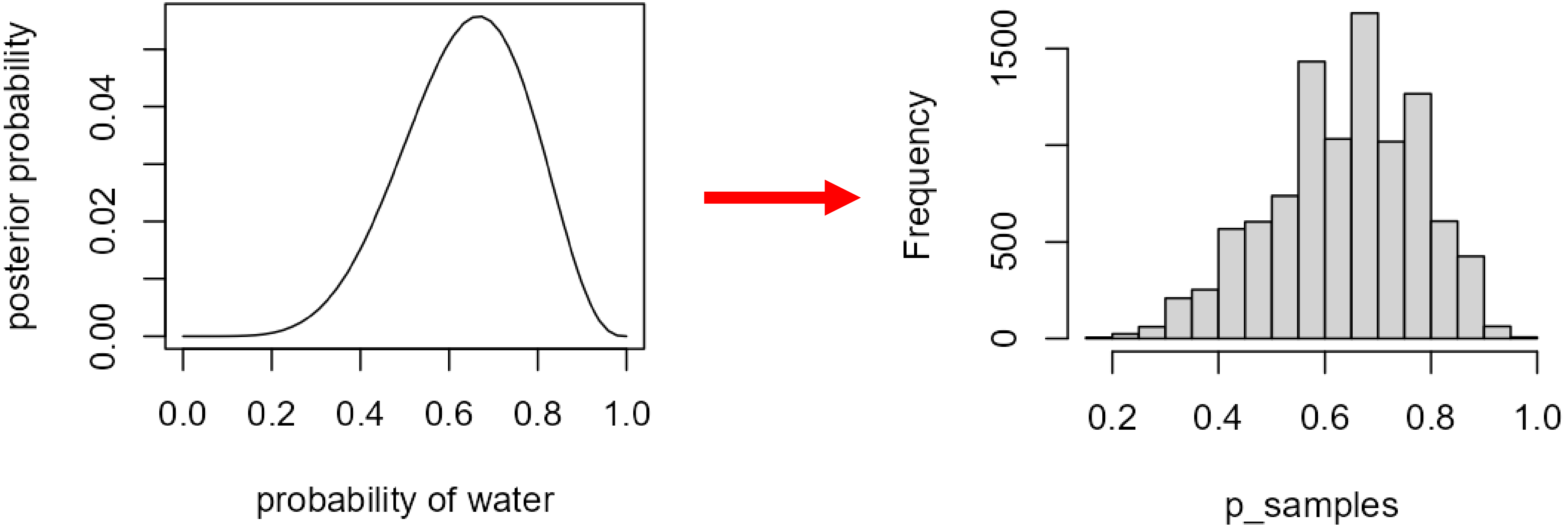
- Globe tossing example: choose a discrete set of p values and compute the posterior probability for each p value

- Can we do something similar with R_0 values instead?
  - Problem: $P(D|R_0)$ values are extremely small
  - Could be numerical issues with trying to sample R_0 values from a grid approximated posterior

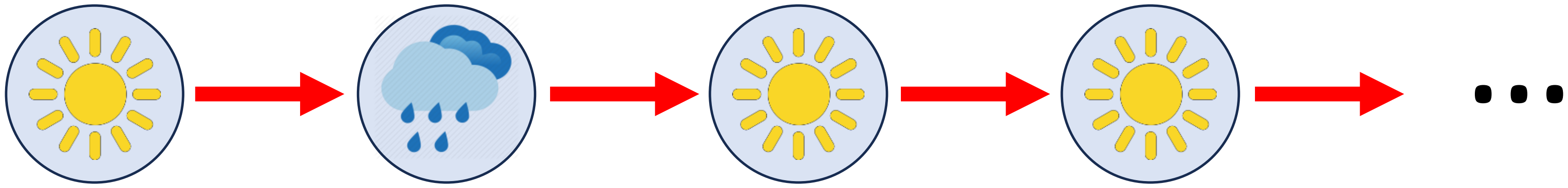
- Grid approximation computationally infeasible with high-dimensional parameter spaces

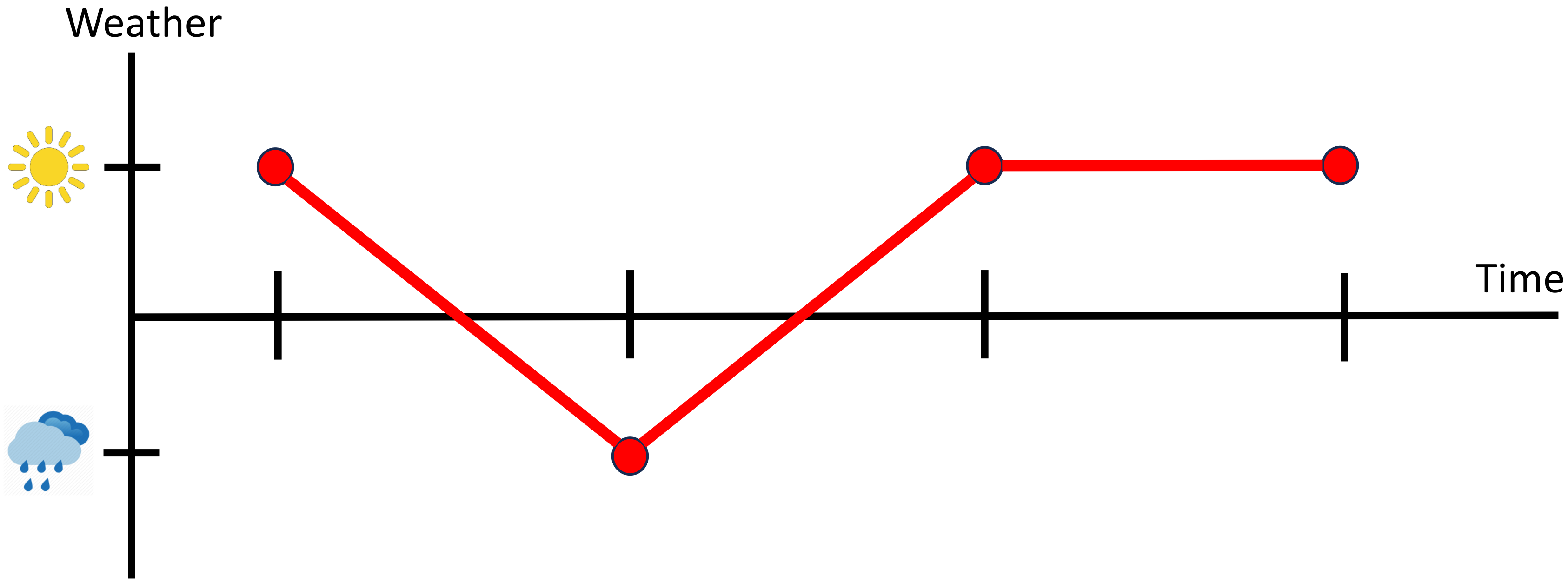
Markov Chain Monte Carlo (MCMC)

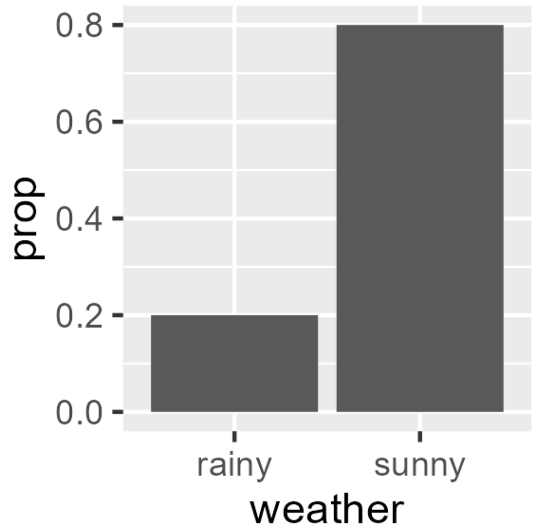
- Algorithm for obtaining **samples** from the posterior distribution without ever computing the actual posterior distribution



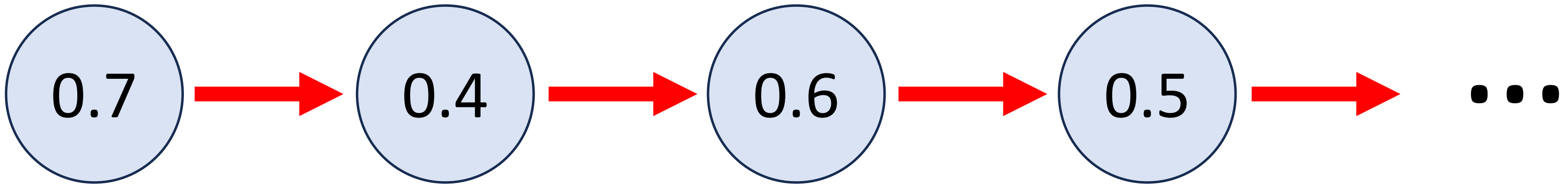
Markov chains (weather space)

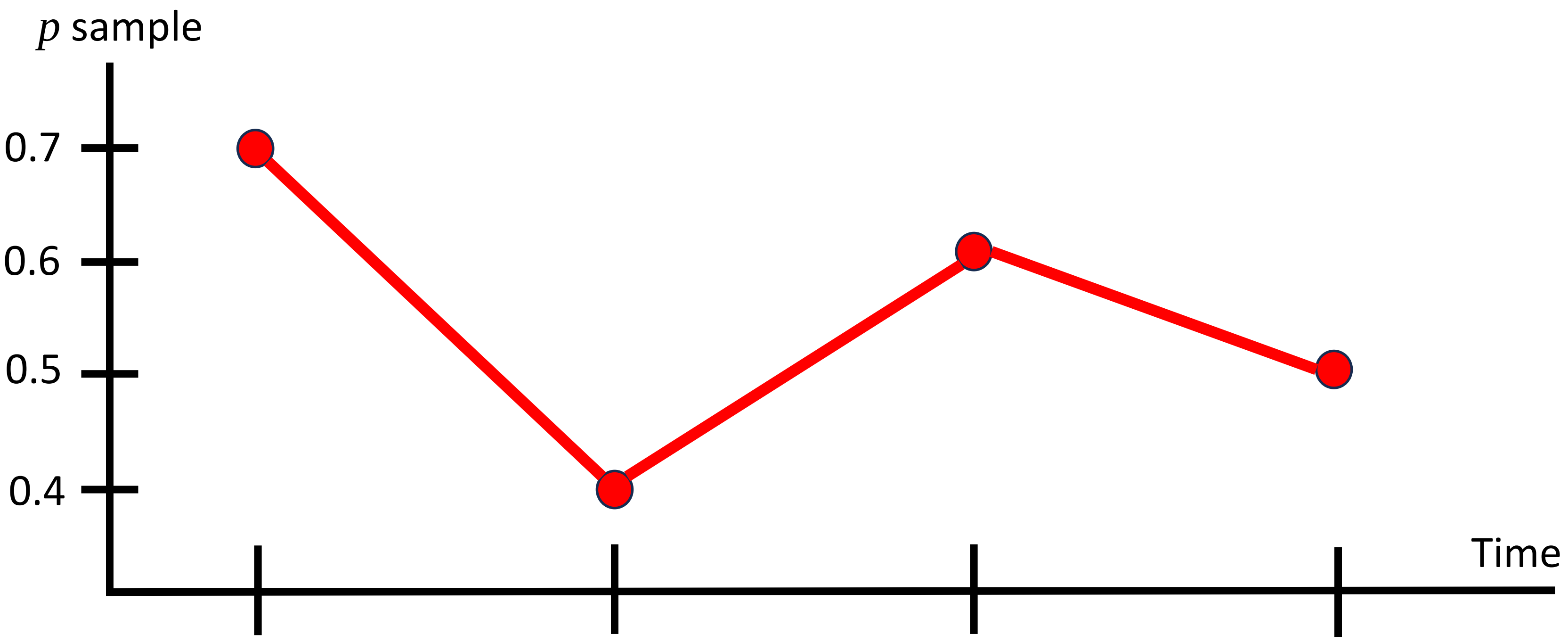


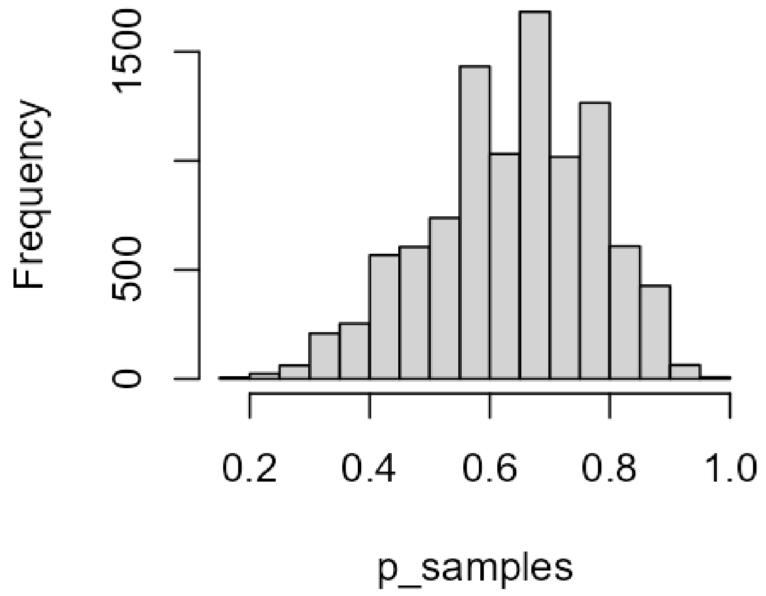




Markov chains (proportion of water *p* space)

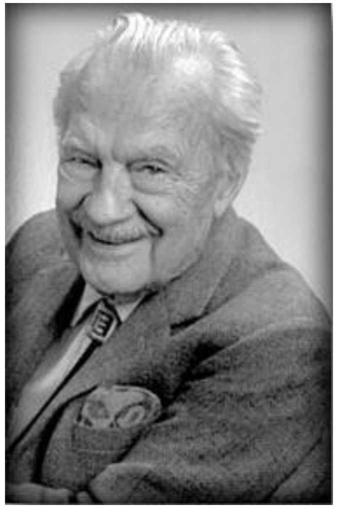






Metropolis-Hastings Algorithm

- Constructs a Markov chain that gives us parameter samples in proportion to their posterior probability

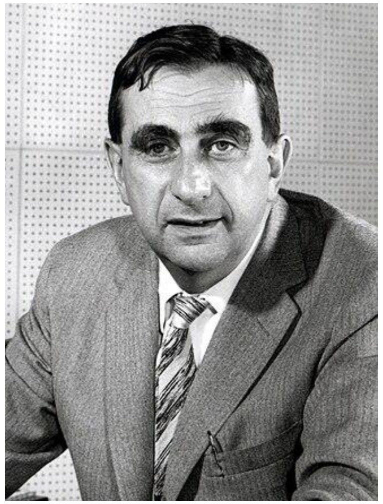


Nicholas
Metropolis

Arianna
Rosenbluth

Marshall
Rosenbluth

Augusta
Teller



Edward
Teller

Wilfred
Hastings

Metropolis Algorithm

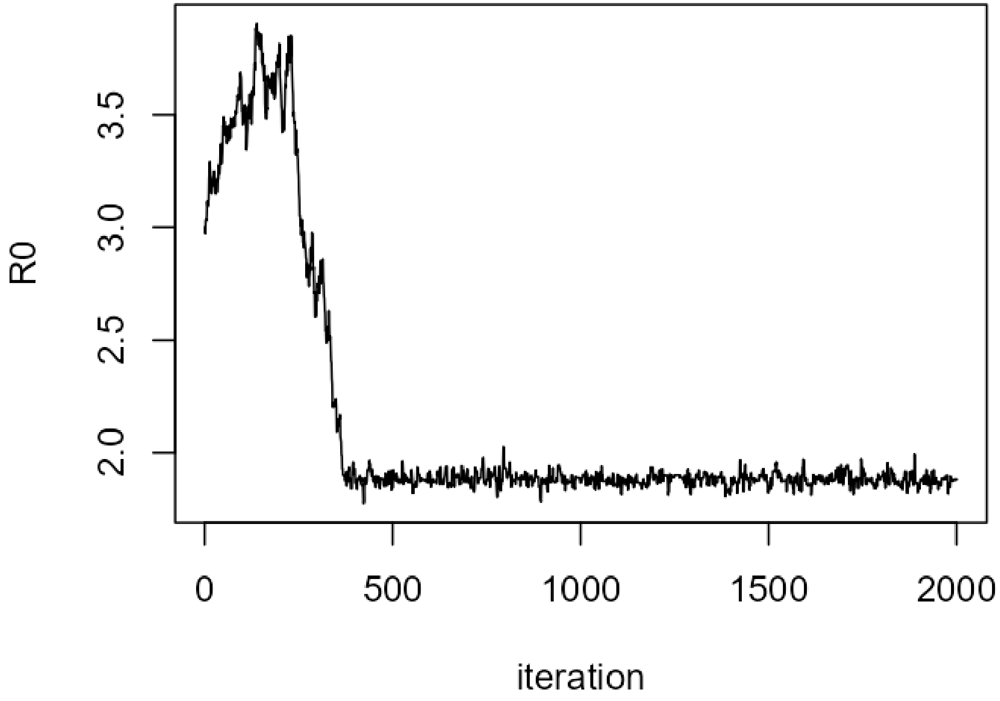
1. Start the chain at a parameter value θ_0

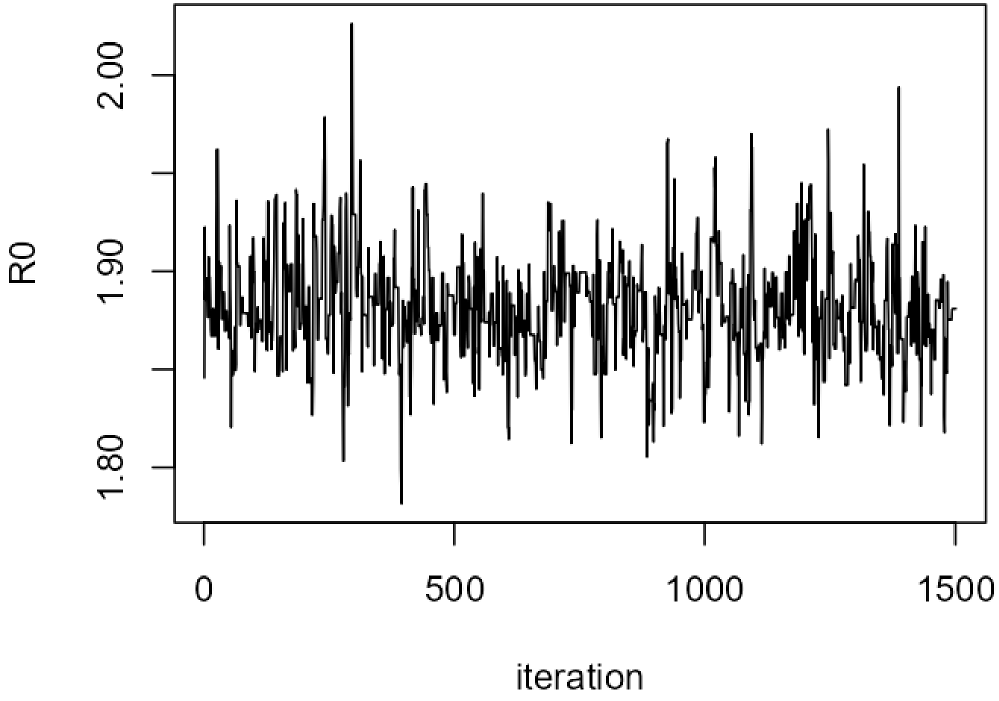
2. Choose a new parameter value (θ^*), based on the current parameter value (θ)

3. If the $P(D|\ \theta^*)P(\theta^*) > P(D|\theta)P(\theta)$, move the chain to θ^*. Otherwise, move the chain to θ^* with probability $P(D|\ \theta^*)P(\theta^*)/P(D|\theta)P(\theta)$

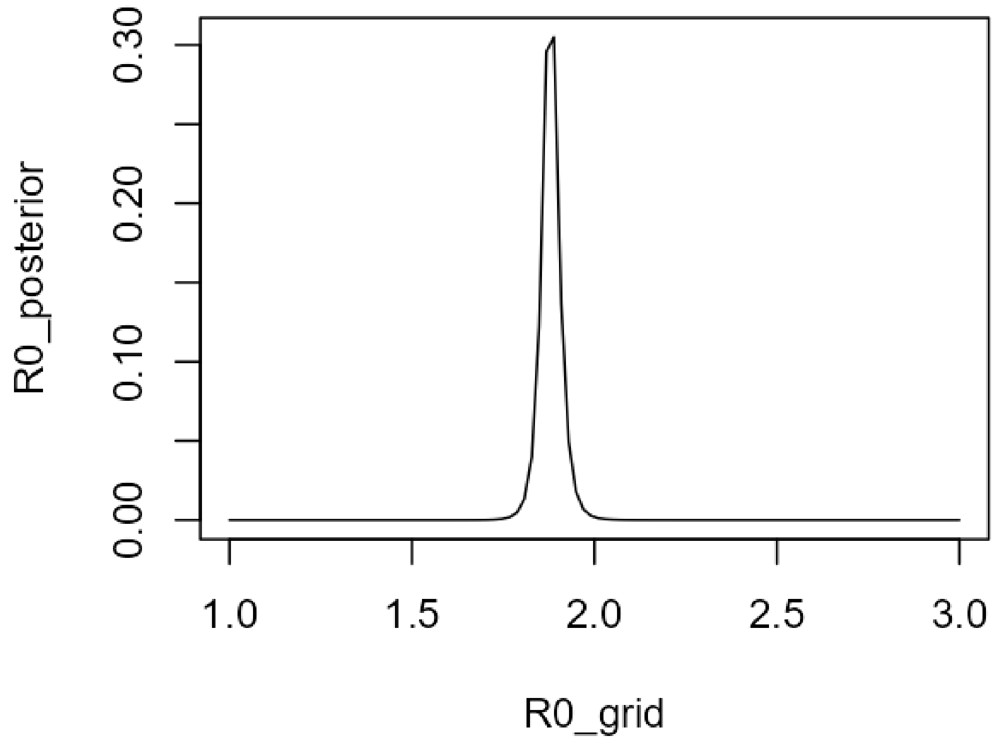
4. Repeat steps 2-3 until the distribution of parameter values converges

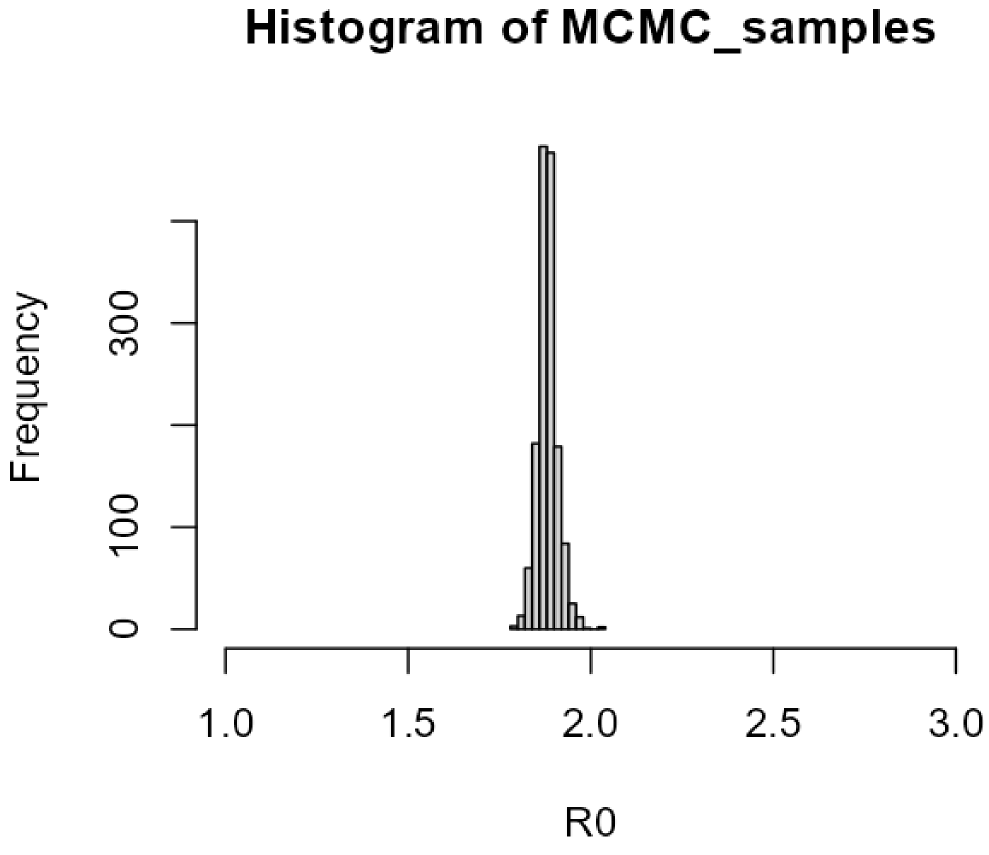
Burn-in samples need to be removed



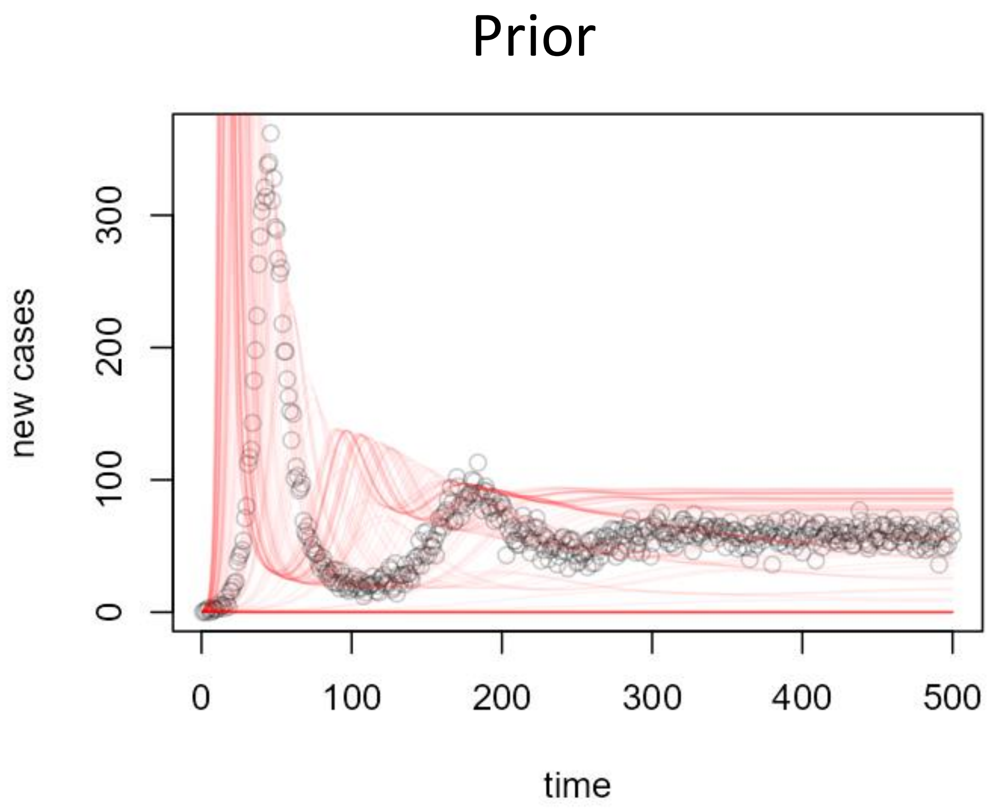


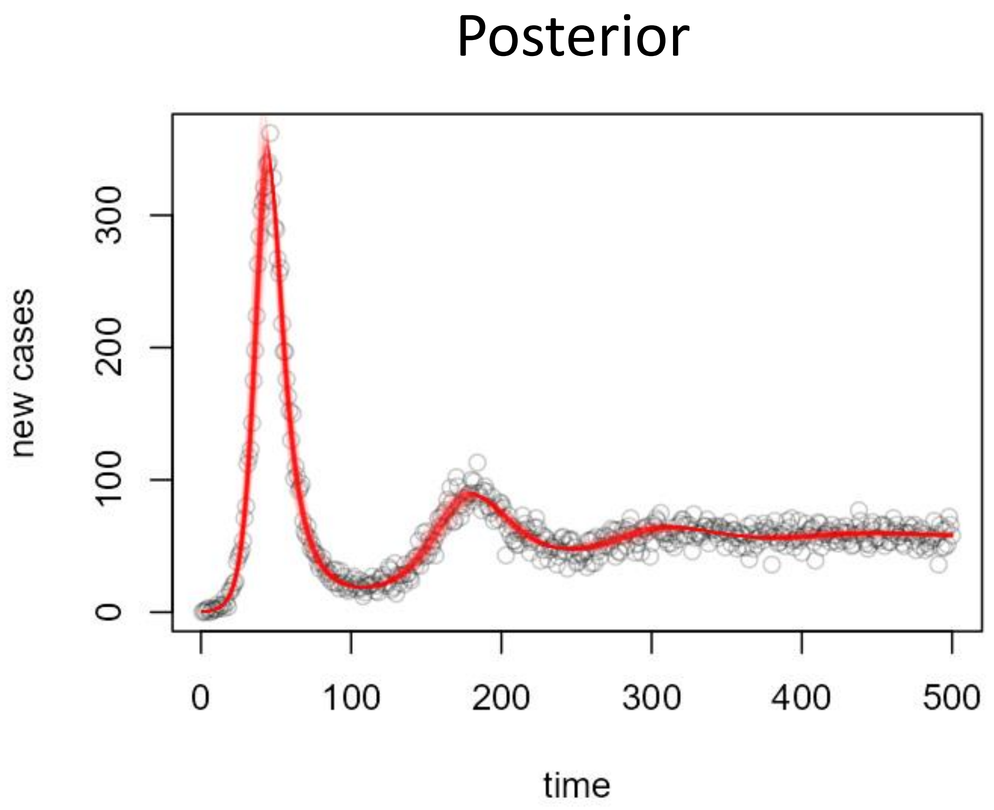
Grid approximation versus MCMC





Prior and posterior mean trajectories





Thanks for listening!