Monte Carlo Markov Chain

Kayla Thomas

galvanize

- Monte Carlo Simulation
- Markov Chains
- Sampling Algorithms

Objectives

- Explain the goal of MCMC
- Describe the output of a MCMC
- Describe the elements of MCMC
    - Monte Carlo simulation
    - Markov Chains
- Describe the sampling process

What is MCMC

MCMC is a method used to estimate a posterior distribution of a parameter of interest by randomly sampling the probabilistic space.

Essentially, we are going to take a new approach to the same models we have learned before, but rather than learning the parameters we are going to learn distributions of the parameters.

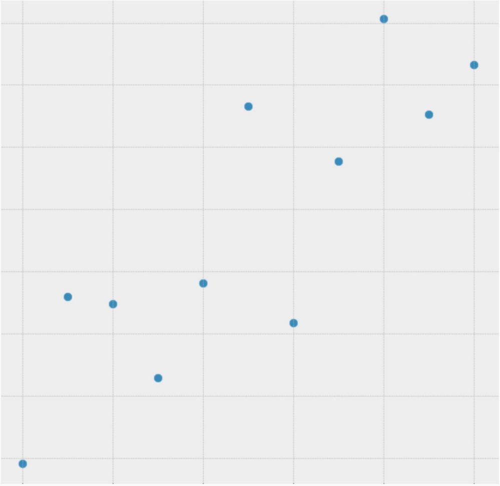
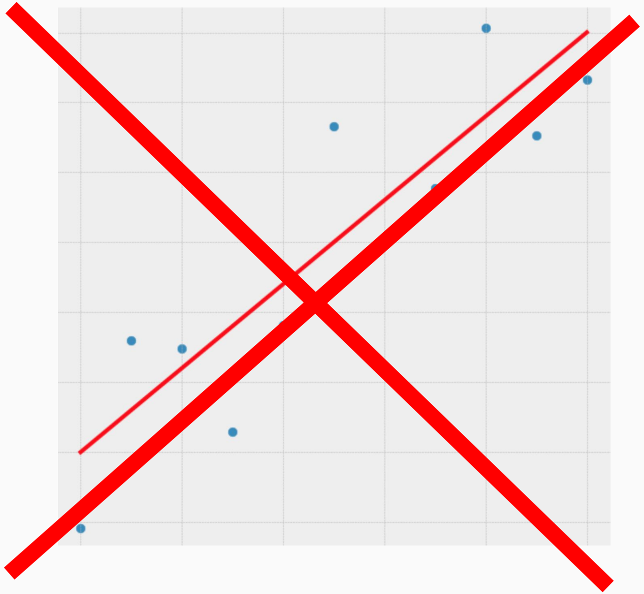
Pros and Cons

Why you might want to use probabilistic programming

1. Customization - We can create models that have built-in hypothesis tests
2. Propagation of uncertainty - There is a degree of belief associated prediction and estimation
3. Intuition - The models are essentially 'white-box' which provides insight into our data

Why you might NOT want use out probabilistic programming

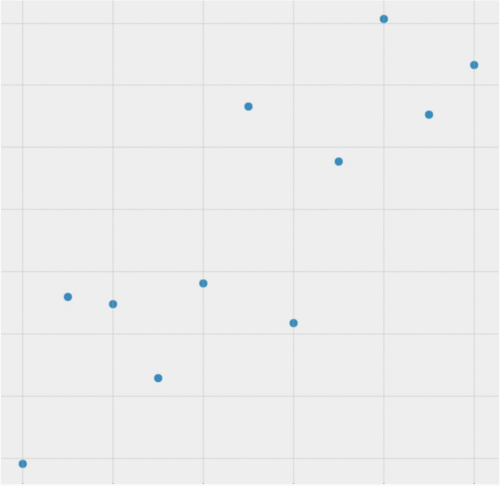
1. Deep dive - Many of the online examples will assume a fairly deep understanding of statistics
2. Overhead - Computational overhead might make it difficult to be production ready
3. Sometimes simple is enough - The ability to customize models in almost a plug-n-play manner has to come with some cost.

Input	Process	Output
	MCMC	

Important distinction

galvanize

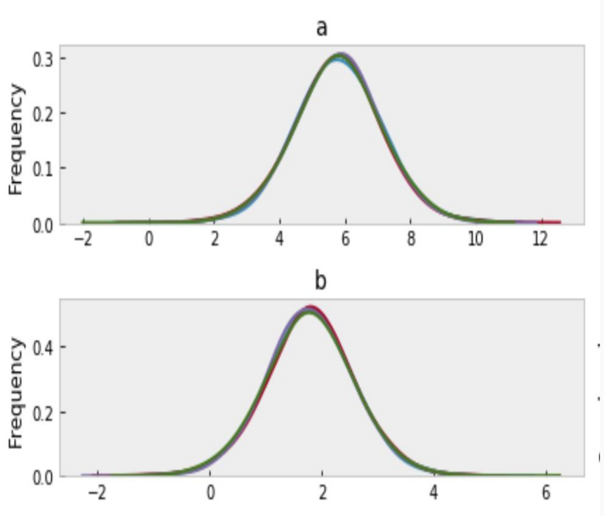
Input



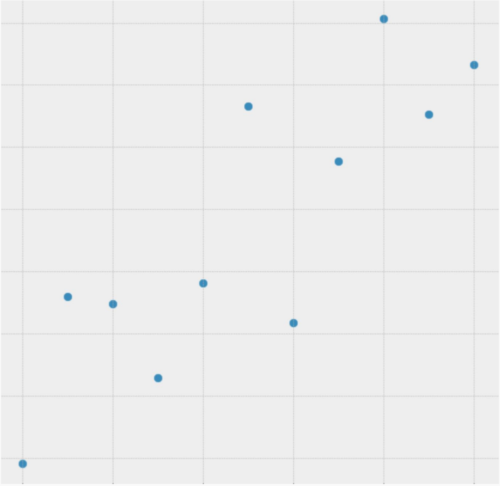
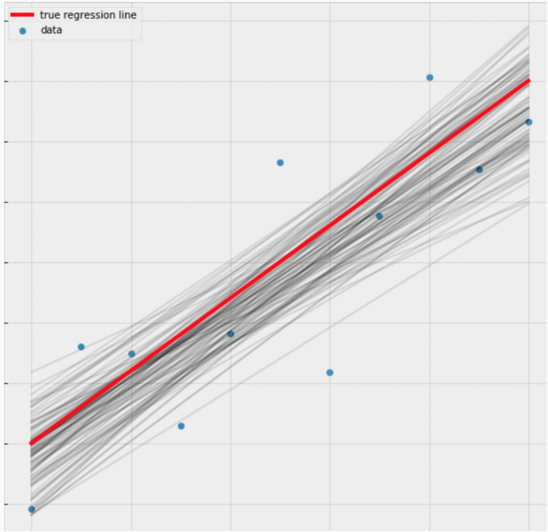
Process

MCMC

Output

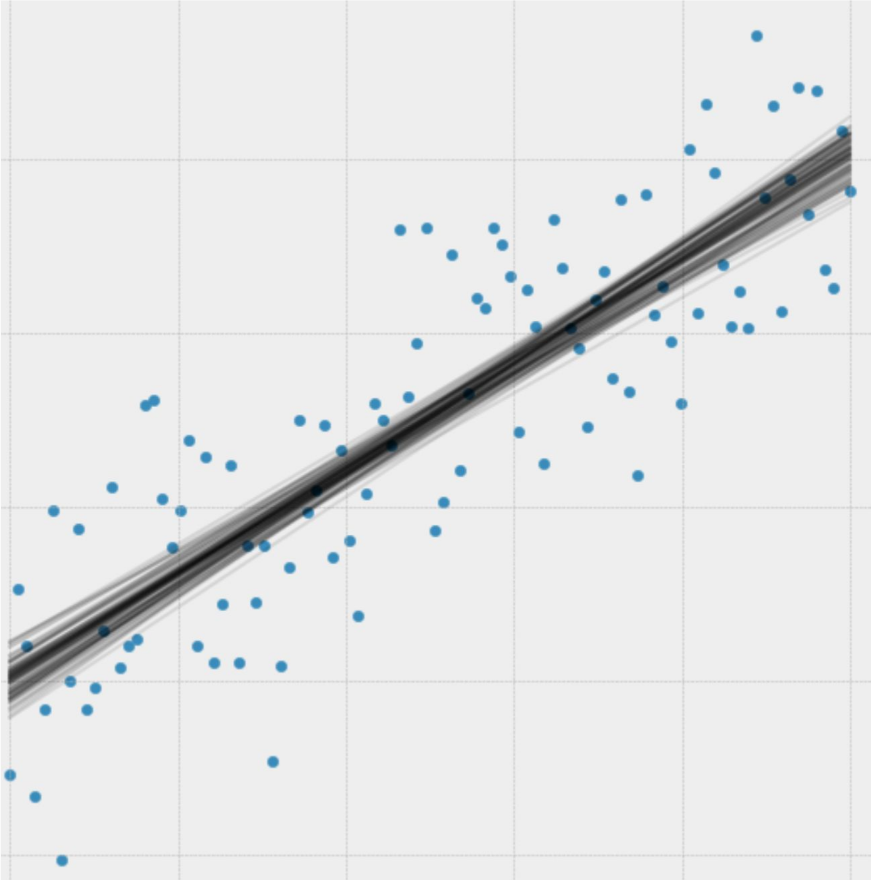


Important distinction

Input	Process	Output
	MCMC	

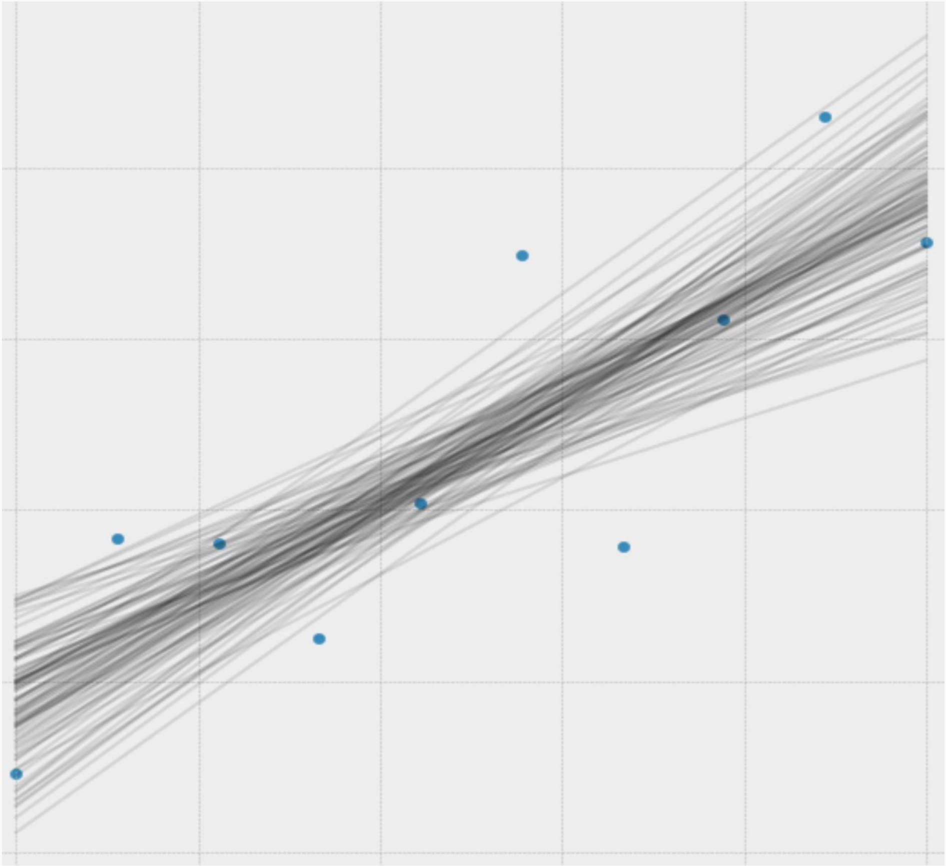
Sampling regression lines

n=100



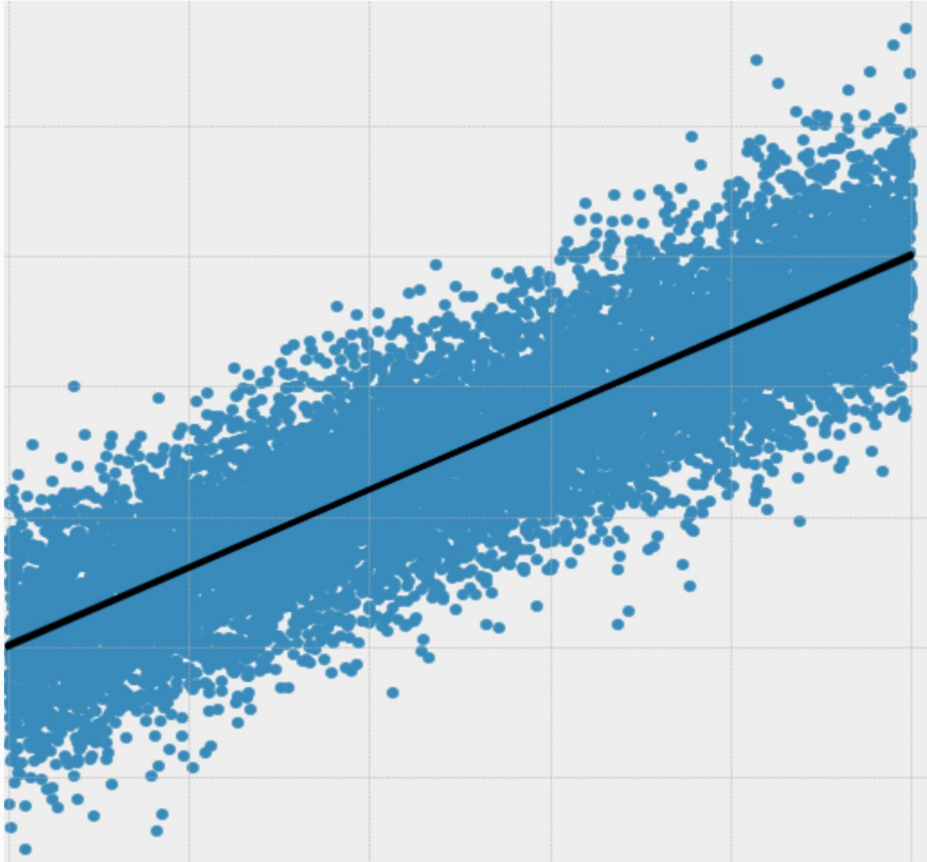
Sampling regression lines

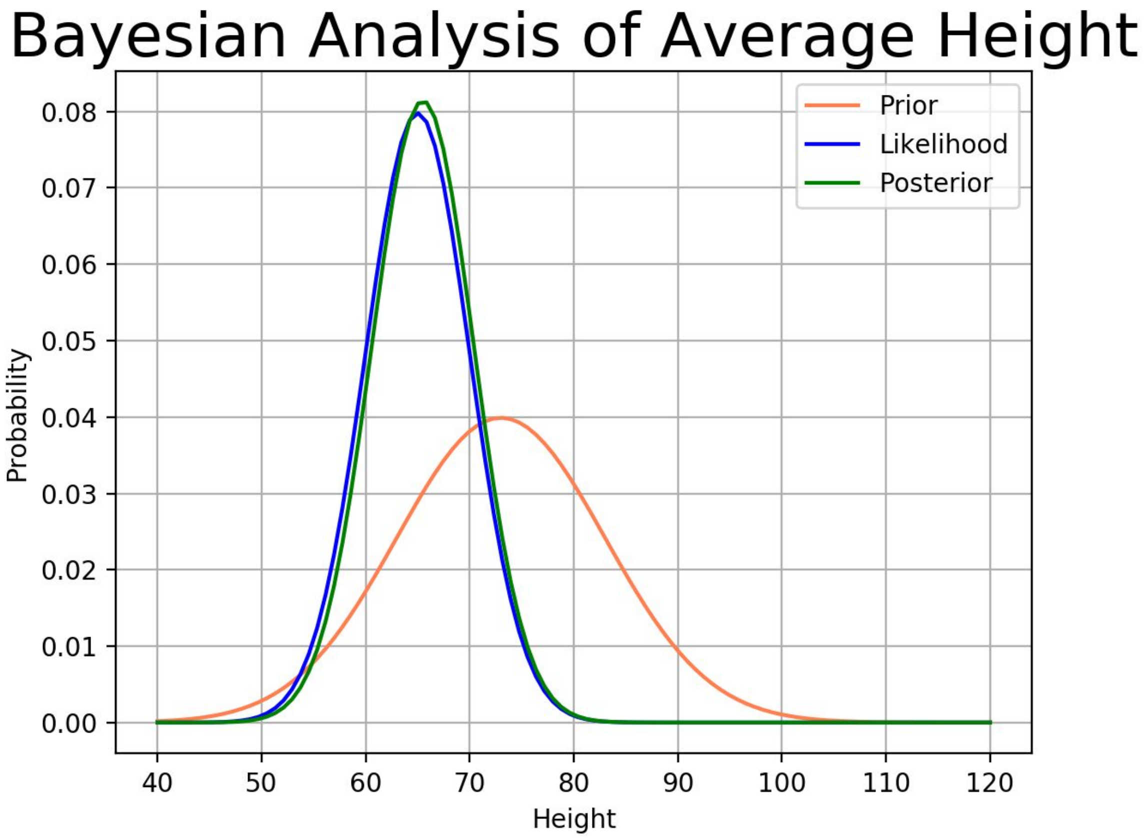
n=10



Sampling regression lines

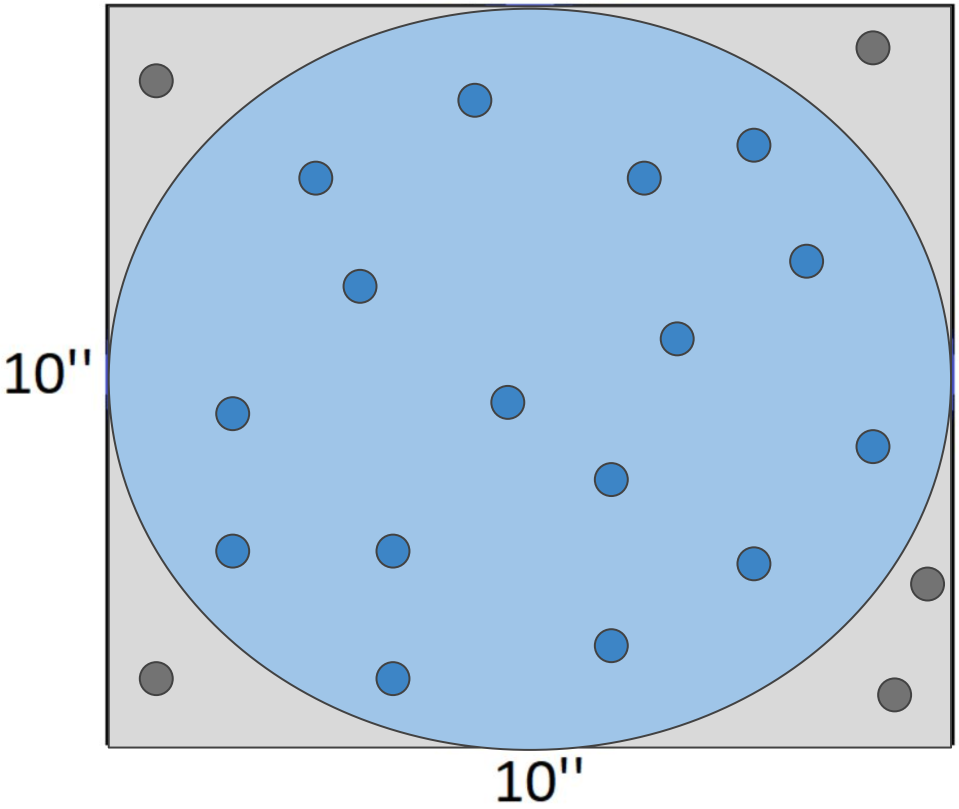
n=10000





What if our prior and likelihood were much uglier?

Area of the Circle= 78.5 inches^2



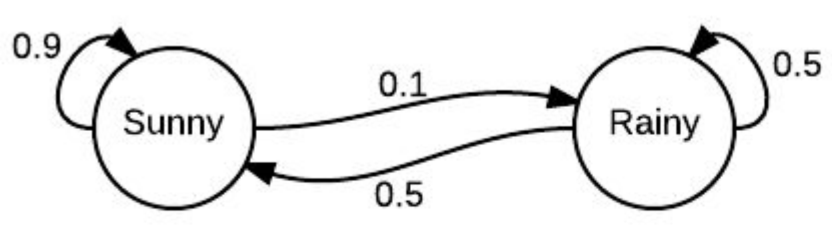
Estimated Area using Monte Carlo simulation:

15 of 20 dots in the circle, about 75 inches^2

Law of large numbers…

In probability theory, the **law of large numbers** (**LLN**) is a theorem that describes the result of performing the same experiment a large number of times. According to the law, the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer as more trials are performed.(Wikipedia)

Pavel Nekrasov vs Andrey Markov



What we need

Some data
A model
A sampler
A likelihood function
An equation for the prior

Let's start with a linear example:

There are many sampling Algorithms, let's use the Metropolis Rule:

To do this we need a Likelihood function, for more one likelihood functions with regression see 1, or 2

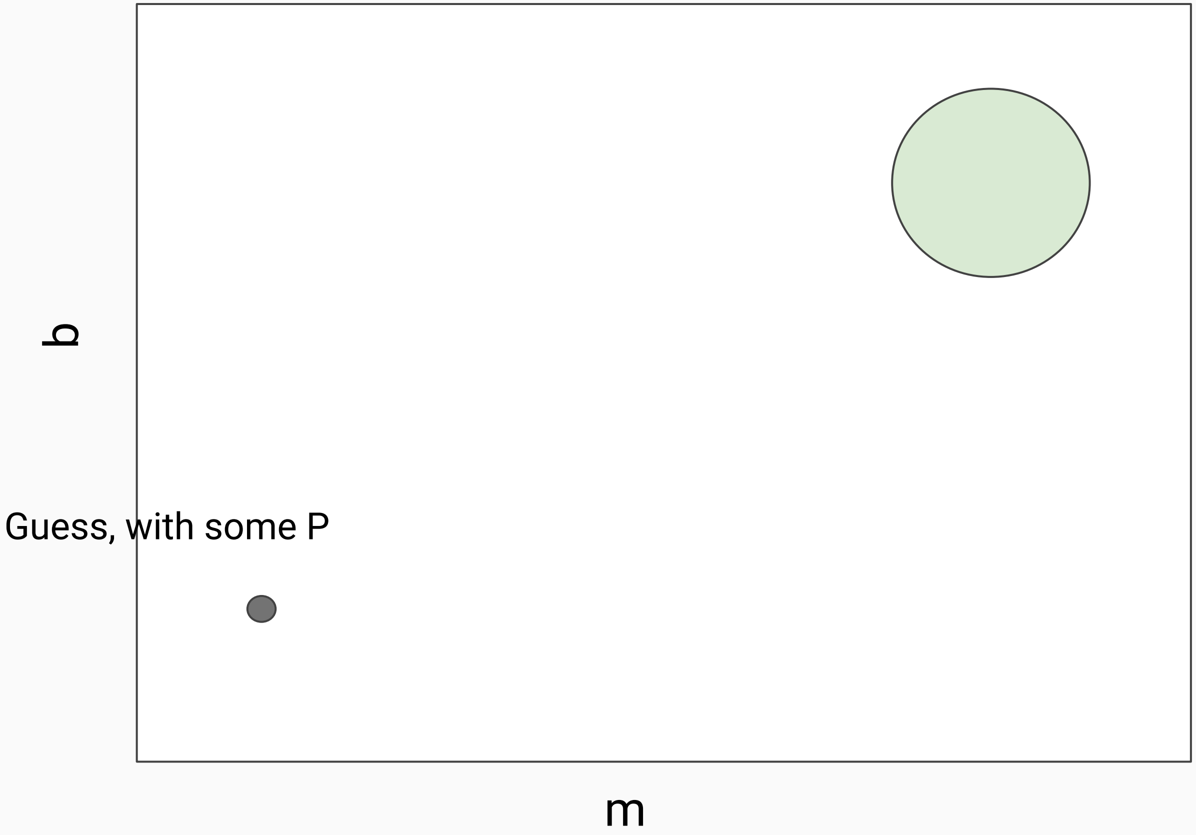
For us we will use:

$$lnL = K - \sum_{i=1}^{n} \frac{|\, y_i - mx_i - b \,|^2}{2\sigma_{yi}^2}$$

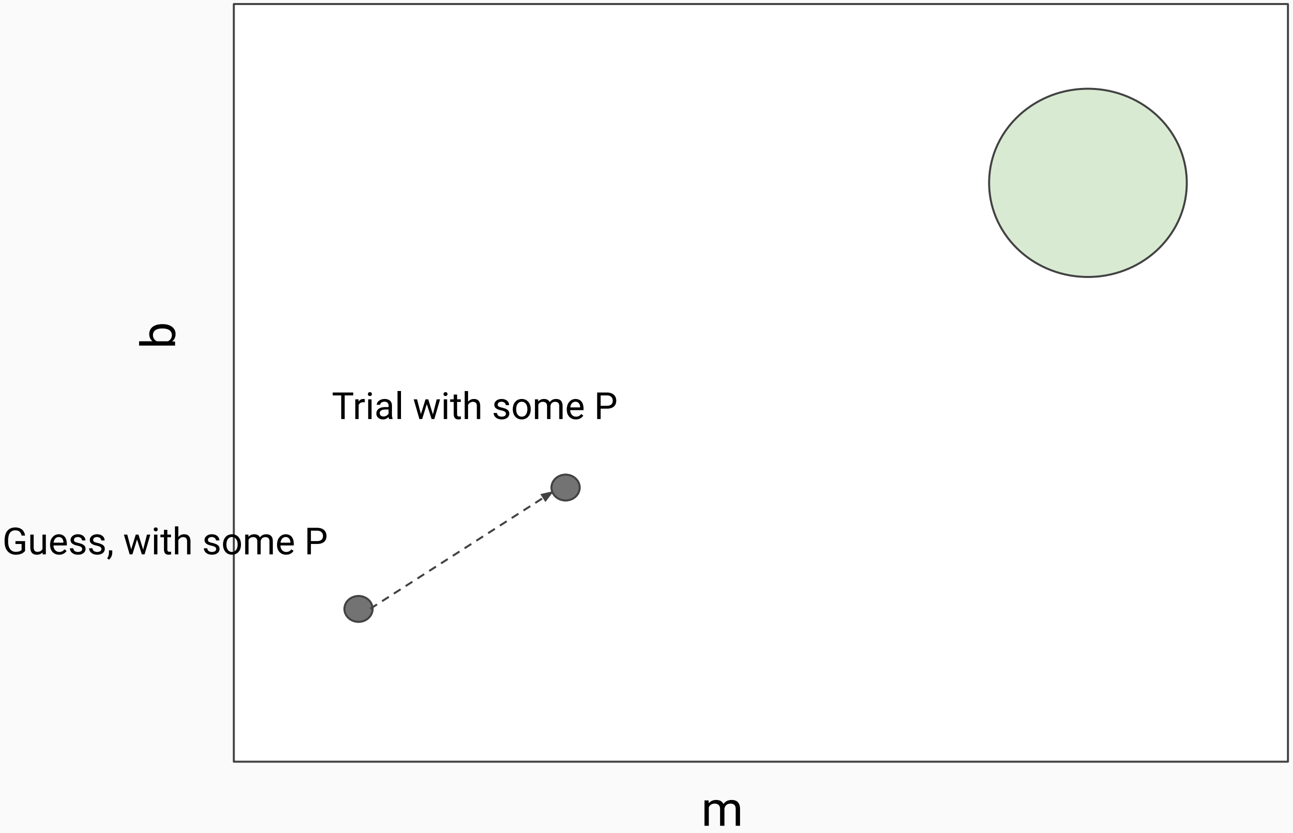
Same as minimizing MSE

First, we will guess for our parameters and calculated the P(yi | xi, sigma, m, b)

Sampling Algorithm - Metropolis Rule



Sampling Algorithm - Metropolis Rule



Metropolis Rule:

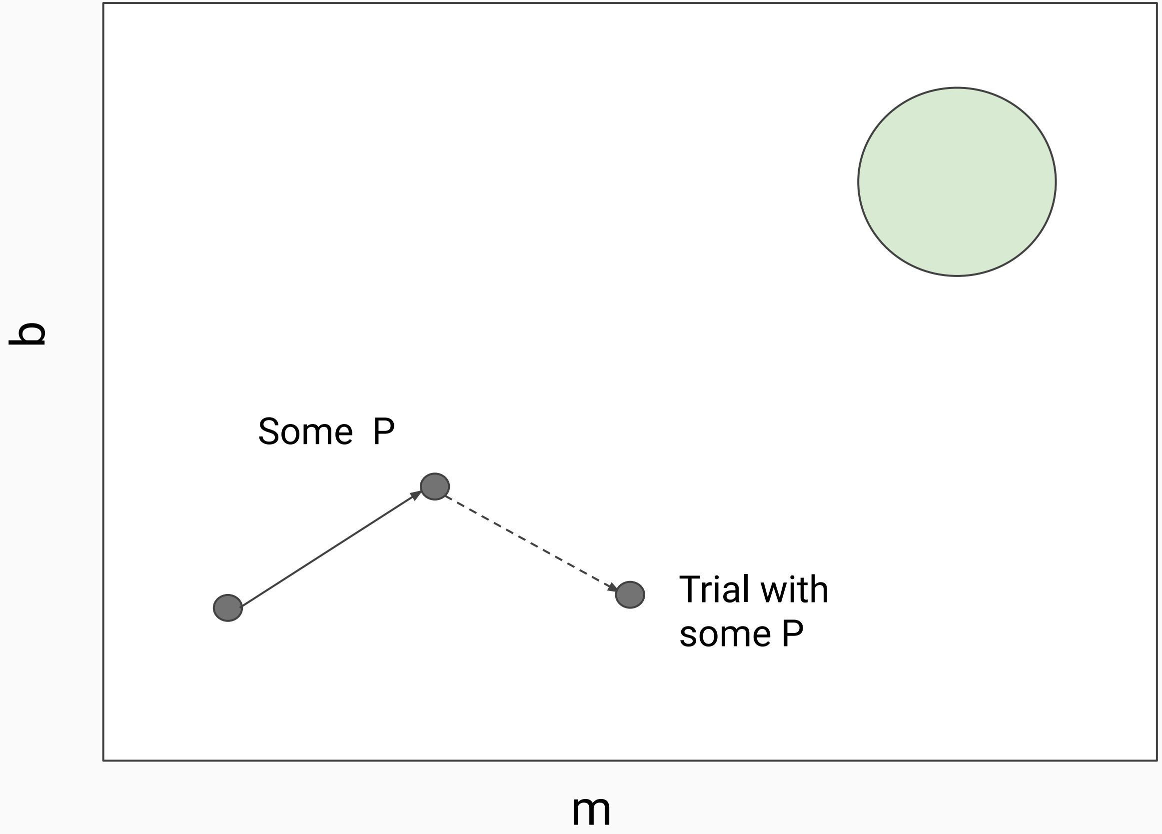
If $P_{Trial} > P_i$

Accept the jump

If $P_{Trial} < P_i$

Accept with Probability P_{Trial} / P_i

Sampling Algorithm - Metropolis Rule



Metropolis Rule:

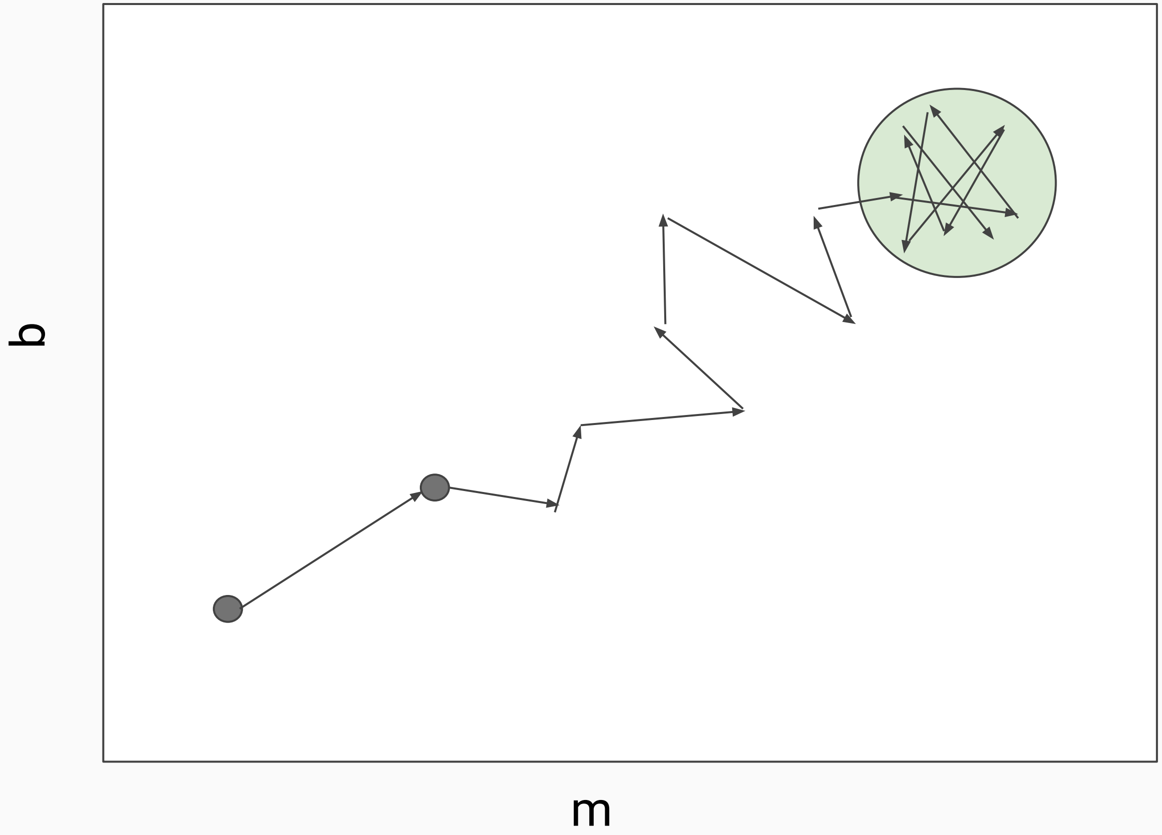
If $P_{Trial} > P_i$

Accept the jump

If $P_{Trial} < P_i$

Accept with Probability P_{Trial} / P_i

Sampling Algorithms - Metropolis Rule



Metropolis Rule:

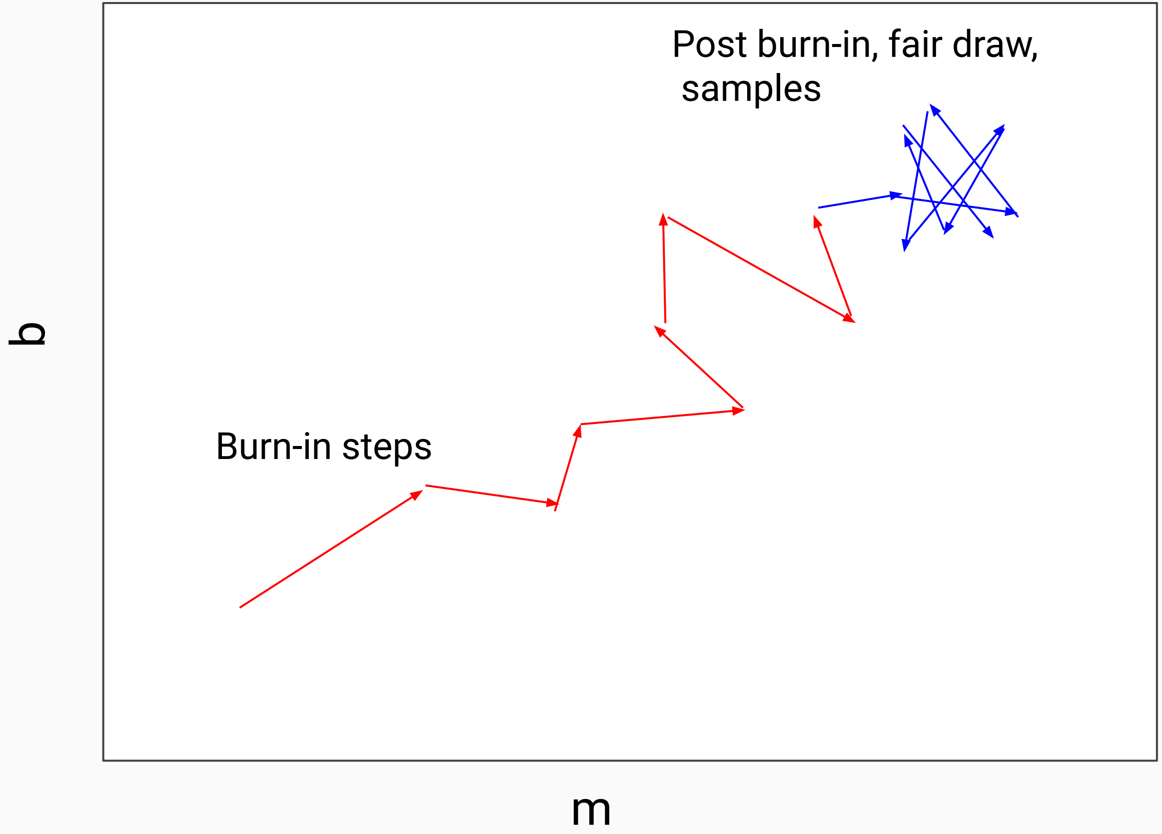
If $P_{Trial} > P_i$

Accept the jump

If $P_{Trial} < P_i$

Accept with Probability P_{Trial} / P_i

Sampling Algorithms - Metropolis Rule



Metropolis Rule:

If $P_{Trial} > P_i$

Accept the jump

If $P_{Trial} < P_i$

Accept with Probability P_{Trial} / P_i

What we need

Some data
A model
A sampler
A likelihood function
An equation for the prior

Linear Example

mcmc_linear_regression.ipynb

Resources

- Adam Richards (Former Denver DSI lead instructor) github
- Pymc3 docs
- David Kipping github
- Why to consider PMC3 blog
- Zero-Math introduction to MCMC blog

Objectives

- Explain the goal of MCMC
- Describe the output of a MCMC
- Describe the elements of MCMC
  - Monte Carlo simulation
  - Markov Chains
- Describe the sampling process