Markov Chain Monte Carlo (MCMC)

Maria

Markov Chain Monte Carlo (MCMC)

- When we talk about Bayesian density and Bayesian inference, it is better if we know Markov Chain Monte Carlo (MCMC).
- MCMC is a different model with Variational Inference (VI)
- The MCMC algorithm aims to generate a sample from a given probability distribution.

Markov Chain Monte Carlo (MCMC)

MARKOV CHAIN MONTE CARLO

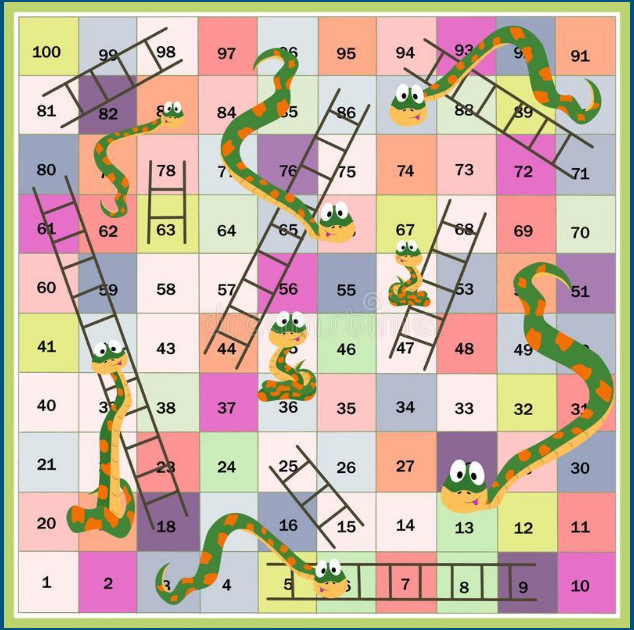
Monte Carlo is part of method's name is sampling purpose. Markov Chain is part come from the way we obtain samples.

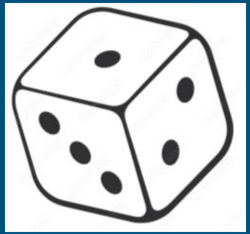
Markov Chain

The conditional probability distribution of the future state of a process depends only on the present state, not on the sequence of events that preceded it.

Markov Chain

Cute and ladders



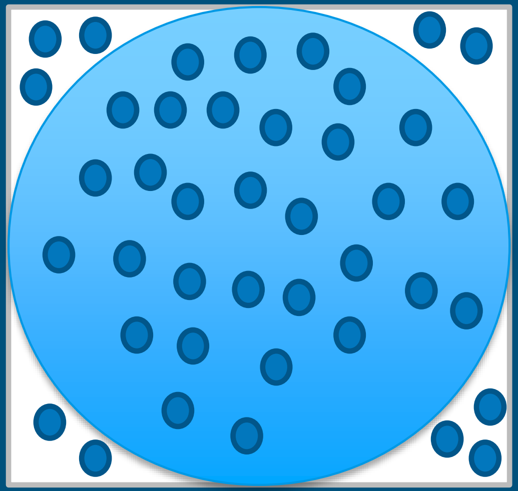


The board consists of 100 numbered squares, with the objective being to land on square 100. The roll of the die determines how many squares the player will advance with equal probability of advancing from 1 to 6 squares. Each move is only determined the player’s present position. However, the board is filled with chutes, which move a player backward if landed on, and ladders, which will move a player forward.

Monte Carlo

- Monte Carlo Experiment or Monte Carlo Simulation.
- An algorithm for obtaining the desired value by performing a simulation involving probabilistic choices.
- Sometimes it's not easy to solve problems using algebra.
- For example is estimating value of π . (common example in Monte Carlo Simulation)

Monte Carlo



Goal : Estimating value of π .

Count :

- Dot inside the circle (C)
- Dot inside the square (S)

$$\frac{C}{S} \approx \frac{\pi \left(\frac{d}{2}\right)^2}{d^2}$$

Markov Chain Monte Carlo (MCMC)

Several MCMC algorithms are commonly used in research or well-known among researchers.

1. Gibbs Sampling Algorithm
2. Metropolis - Hasting Algorithm

Gibbs Sampling

- The method to generate a sequence of samples from the joint probability distribution of two or more random variables.
- This method always accepts all proposals.

initialize Y^0, X^0

for j = 1, 2, 3,… do

  sample $X^j \sim p(X|Y^{j-1})$

  sample $Y^j \sim p(Y|X^j)$

end for

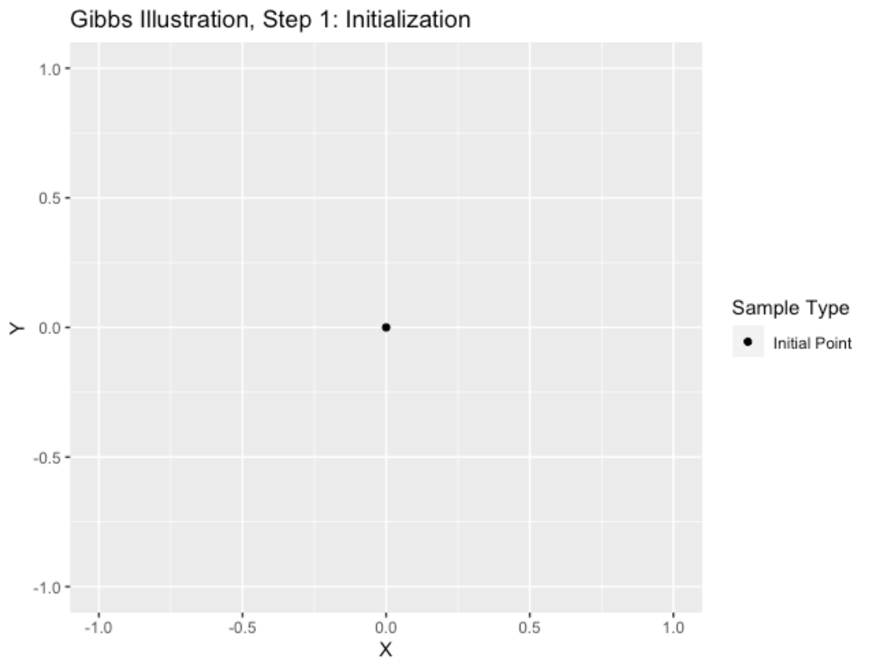
Gibbs Sampling

- Gibbs sampling for a Bivariate Normal target distribution with correlation ρ :
  1. Initialize : $\left(x_0, y_0 \; \coloneqq (0,0)\right)$ and set $t \coloneqq 0$
  2. Draw x_t from the conditional distribution $X_t | \, (Y_t - 1 = \; y_t - 1) \sim N(\rho y_t - 1, 1 - \rho^2)$
  3. Draw y_t from the conditional distribution $Y_t | (X_t = x_t) \sim N(\rho x_t, 1 - \rho^2)$
  4. Increment $t \coloneqq t + 1$
  5. Return to step 2

Dobrow, R.P. 2016. *Introduction to Stochastic Processes with R* (Wiley).

Gibbs Sampling

- Set first iteration of our Gibbs sampler with ρ equal to 0.9.
- Then set (x_o, y_o) to (0,0).



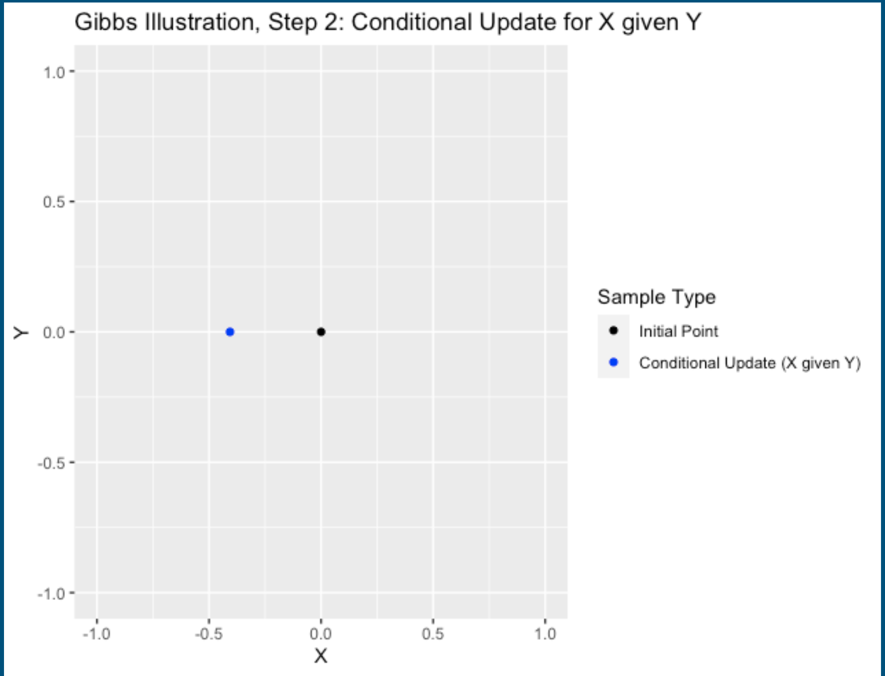
https://towardsdatascience.com/gibbs-sampling-explained-b271f332ed8d

Gibbs Sampling

- We draw from the conditional distribution of X given Y equal to 0.
- Condition X given Y

$$X_1 \mid (Y_0 = 0) \sim N\,(0 \cdot \ \rho, 1 - \rho^2)$$

- The result is – 0.4

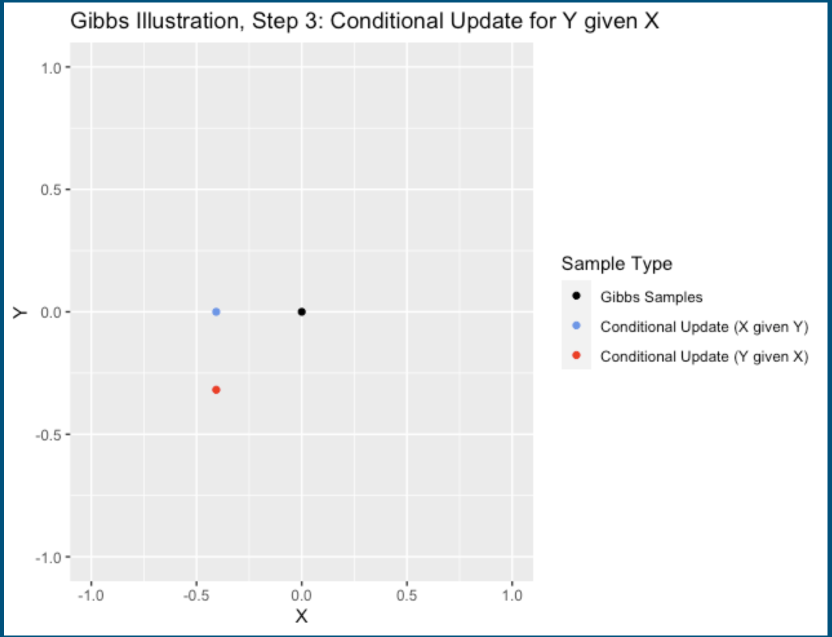


Gibbs Sampling

- We draw from the conditional distribution of Y given X equal to -0.4.
- Condition Y given X:

$$Y_1|(X_1 = -0.4) \sim N(-0.4.\rho, 1 - \rho^2)$$

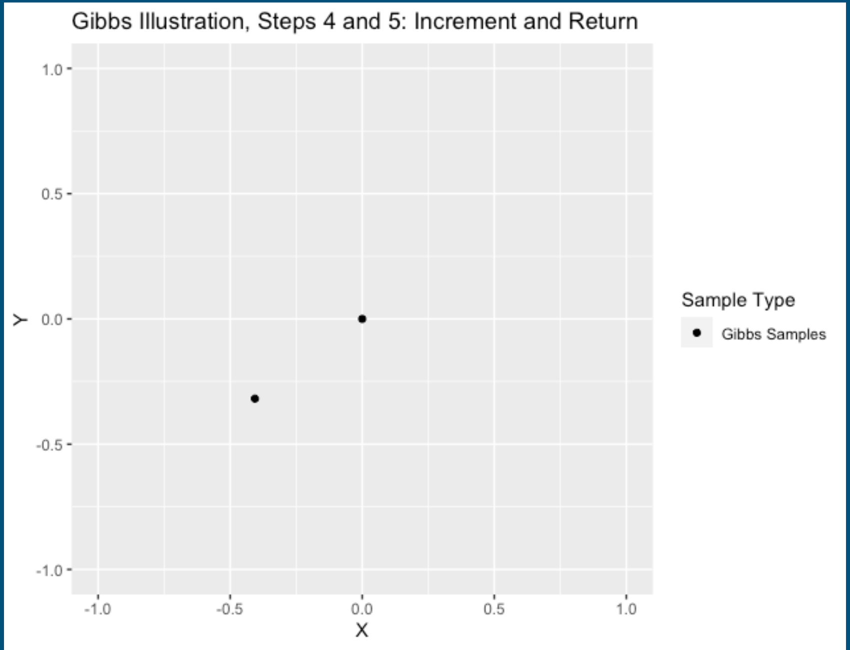
- The result is - 0.32



https://towardsdatascience.com/gibbs-sampling-explained-b271f332ed8d

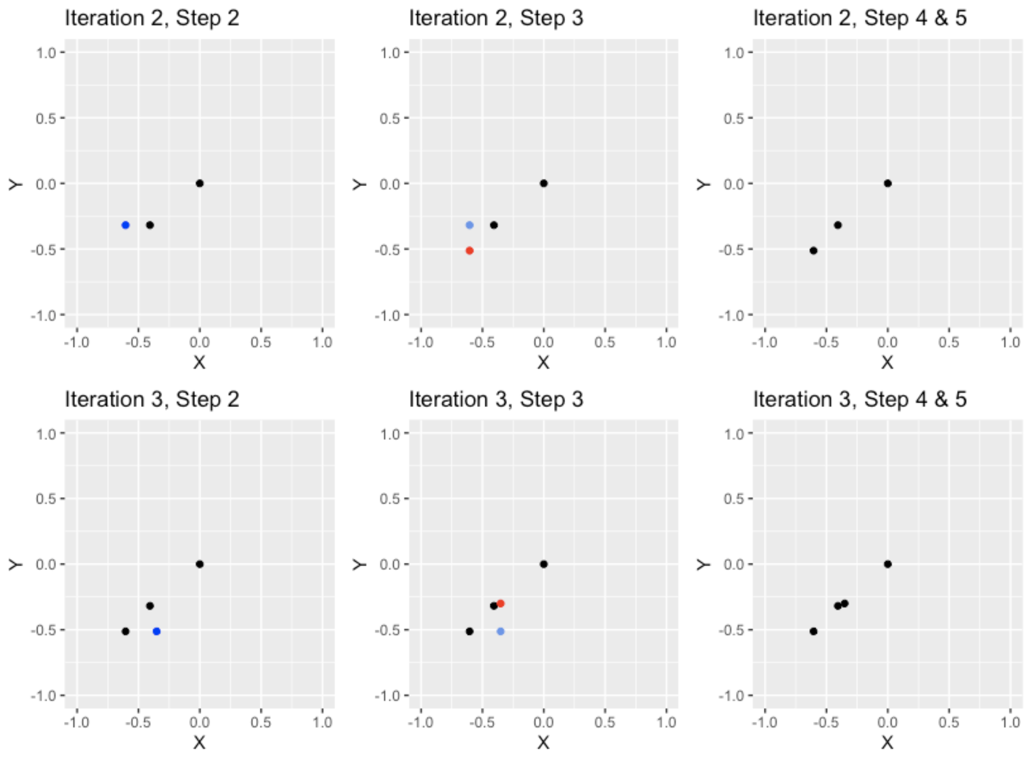
Gibbs Sampling

- The final point for this iteration of the Gibbs sampler is (-0.4, -0.32)



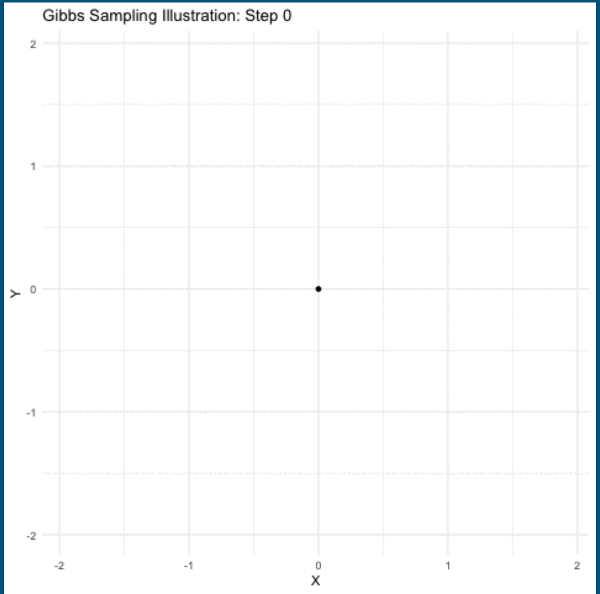
https://towardsdatascience.com/gibbs-sampling-explained-b271f332ed8d

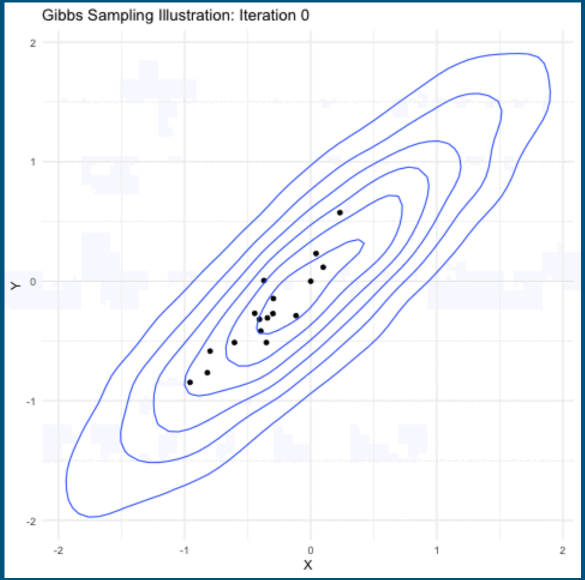
Gibbs Sampling



https://towardsdatascience.com/gibbs-sampling-explained-b271f332ed8d

Gibbs Sampling





https://towardsdatascience.com/gibbs-sampling-explained-b271f332ed8d

Metropolis - Hasting Algorithm

- Instance of Markov Chain Monte Carlo algorithm is Metropolis – Hasting Algorithm.
- This algorithm also called "random walk", where the distribution is repeatedly sampled in small steps; independent of the previous step, and so is no memory (memoryless).

Metropolis - Hasting Algorithm

A politician is campaigning in 7 districts, one adjacent to the other. She wants to spend time in each district, but due to financial constraints, would like to spend time in each district proportional to the number of likely voters in that district. The only information available is the number of voters in the district she is currently in, and in those that are directly adjacent to it on either side. Each day, she must decide whether to campaign in the same district, move to the adjacent eastern district, or move to the adjacent western.

https://www.publichealth.columbia.edu/research/population-health-methods/markov-chain-monte-carlo

Metropolis - Hasting Algorithm

On any given day, here's how the decision is made whether to move or not:

- Flip a coin. Heads to move east, tails to move west.
- If the district indicated by the coin (east or west) has more voters than the present district, move there.
- If the district indicated by the coin has fewer likely voters, make the decision based on a probability calculation

Metropolis - Hasting Algorithm

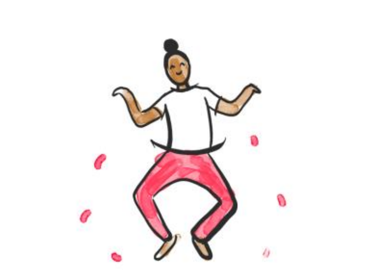
- If the district indicated by the coin has fewer likely voters, make the decision based on a probability calculation:
  - Calculate the probability of moving as the ratio of the number of likely voters in the proposed district, to the number of voters in the current district: **P[move] = voters in indicated district/voters in present district**
  - Take a random sample between 0 and 1.
  - If the value of the random sample is between 0 and the probability of moving, move. Otherwise, stay put.

Metropolis - Hasting Algorithm

1

She is in district 4.

The prosed move is to district 5



Metropolis - Hasting Algorithm

2

The prosed move is to district 6

Metropolis - Hasting Algorithm

3

The proposed move is to district 7

Metropolis - Hasting Algorithm

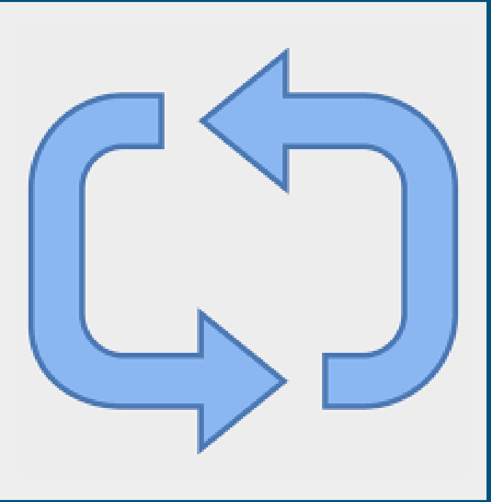
4

The proposed move is to district 6

Metropolis - Hasting Algorithm

Base on the decision to move on the probability criterion of 6/7. Draw a random sample between 0 and 1, if the value is between 0 and 6/7, move. Otherwise, stay in district 7.

Metropolis - Hasting Algorithm



Perform this procedure many times.


Thank You