# Hawaiian Shirt: No!

Told Byron: No.

Date: Tuesday, February 12, 2019

## Announcements

- HW #2 due Thursday @ midnight
- Read Ch. 3 of Probabilistic Programming for Hackers by Thursday
- Almost done with grading... will have key on Thursday

[ Slide 19 – Markov Chain Simulation ]

$$\mathbb{P} = \begin{array}{cc} & \text{to} \\ \text{from} & \begin{pmatrix} & & \\ & P_{ij} & \\ & & \end{pmatrix} \end{array}$$

$P_{ij} = \mathbb{P}$(moving from state $i$ at time $t$ to state $j$ at time $t+1$).

**Example 1:**

state space

**Question 2:** $S = \{1, 2\}$

where 1 = no rain
2 = rain

$$\mathbb{P} = \begin{array}{ccc} & 1 & 2 \\ 1 & 1-\beta & \beta \\ 2 & 1-\alpha & \alpha \end{array}$$

Add rows to 1
So yes, this is MC!

## Example 2:

$S = \{1, 2, 3\}$

where 1 = Cheerful
2 = So-so
3 = Glum

$$\mathbb{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{pmatrix} \end{array}$$

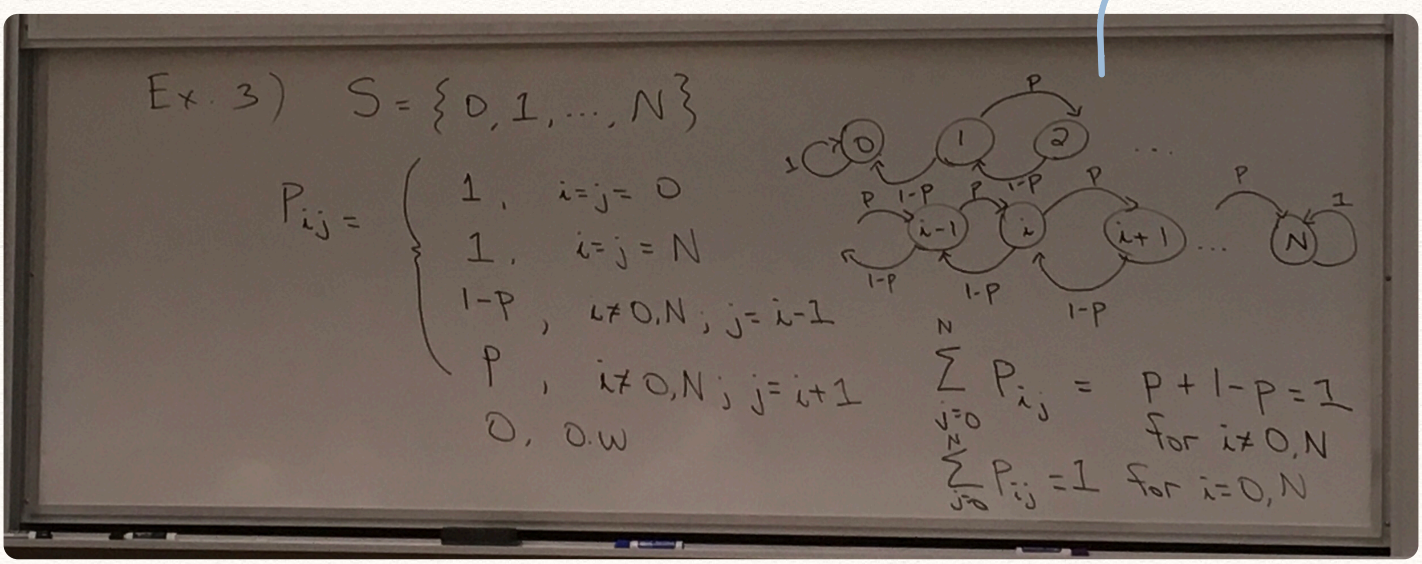
} Rows add to 1, so yes, $\{X_t, t \geq 0\}$ is a MC.

## Example 3:

$S = \{0, 1, \dots, N\}$

$$P_{ij} = \begin{cases} 1 & \text{if } i=j=0 \quad \text{(Goes broke)} \\ 1 & \text{if } i=j=N \quad \text{(Attains fortune of \$N)} \\ 1-p & \text{if } i \neq 0, N;\ j = i-1 \quad \text{(Loses)} \\ p & \text{if } i \neq 0, N;\ j = i+1 \quad \text{(Wins)} \\ 0 & \text{otherwise} \end{cases}$$

State transition diagram



# Stationary Distribution of a DTMC

$\pi_j := \mathbb{P}(X_t = j)$

= Long run probability that your stochastic process visits state $j$.

It can be calculated using the equation:

$$\pi_j = \sum_{i \in S} \pi_i P_{ij}$$

This means we can solve $\pi = \pi \mathcal{P}$ to get the stationary distribution $\pi$.

[Markov Chain Monte Carlo]

# MCMC Basics

* We will start w/ some proposal $g_1(\theta)$
* Subsequently update at each step $t$ (usage $g_t(\theta)$)
  - ↳ The way to do <u>this</u> is to simulate from a transition probability distribution $g_t(\theta) = T_t(\theta^t \mid \theta^{t-1})$ so that the standard distribution of the Markov Chain with transition probability matrix $\Pi_t$ is exactly $\mathbb{P}(\theta \mid y)$.
* Key take away: If you sample long enough, eventually we'll sample from the posterior! :)
* An (unfortunate) take away is that the first few (maybe thousand) samples could be complete garbage.
* This is taken into account with a <u>BurnIn</u> parameter, which specifies how many samples should be tossed out.

* NOW our goal is to come up with smart $T_t$'s so that we get $\mathbb{P}(\theta|y)$ as our stationary distribution.

**Three main algorithms to do this:**

1) Gibbs Sampler

- Most efficient
- !⇒ Requires knowing the conditional distribution of $\theta^t|\theta^{t-1}$ which is rarely ever known :(
- will likely never use

2) Metropolis Algorithm

- 2nd most efficient
- Requires a symmetric Proposal distribution, which we can always come up with :)
- will Probably use the most often

3) Metropolis-Hastings Algorithm

- slowest
- requires no information
- generalizes the Metropolis algorithm
- It can readily be used, but typically stick w/ Metropolis.

$$\begin{pmatrix} \theta_1 \\ \vdots \\ \theta_d \end{pmatrix} \to \mathbb{P}(\theta_1, \ldots, \theta_d | y)$$

there is dependence among parameters

**A valid Symmetric proposal:**

$J(\theta^t | \theta^{t-1})$ = Density of $N(\theta^{t-1}, 1)$

Any Normal distribution w/ mean $\theta^{t-1}$ works! (*)

$$f(\theta a|\theta b) = \frac{1}{\sqrt{2\pi}} e^{\frac{-(\theta_a-\theta_b)^2}{2}} \stackrel{?}{=} f(\theta b|\theta a)$$