

Date: Tuesday, January 31, 2019

## Announcements

- Next week:
  1. Exam on Thursday @ 3:30 - 5:30
    - Multiple choice or fill in the blank
    - Covers Ch 1+2 in book + math insights
    - Distributions:  $\text{Poi}(\lambda)$ ,  $\text{Exp}(\alpha)$ ,  $\text{U}(a,b)$ ,  $\text{DiscUnif}(a,b)$   
 $\text{Normal}(\mu, \sigma^2)$ ,  $\text{Bin}(n,p)$
  2. Have class T, Th, F  
Friday times: 10:00 - 11:55  
2:00 - 3:55
- 3. what are they used for? Soon...
- HW due tonight



3 people have the same birthday in session 1!

Aug 28 : Mawu, Tiangi , Jialiang  
(PS: James' birthday is on Aug 25)

## Chapter 2:

Key concepts: 1) Working w/ PyMC

- Variable relationships
- Stochastic vs. deterministic
- Specifying a Bayesian model
- Posterior inference via simulation

2) Applications of Bayesian Modeling

- Bayesian A/B testing
- Bayesian truth serum (aka "Privacy Algorithm")
- The Challenger

### 3) Goodness of fit

- A means to validate a model via simulation

Distributions discussed :  
-  $\text{Bin}(n, p)$   
-  $\text{Normal}(\mu, \sigma^2)$

## 1) Working with PyMC

Example : Text messages

Data:  $C_i \rightarrow$  count of incoming messages on day  $i$

$$C_i \sim \text{Po}(\lambda)$$

$\lambda$  mean count  
(but also the variance, specific to the distribution).

Prior: DO we know anything about  $\lambda$ ?

(Bayesian way of thinking) : Yes!

(Frequentist way of thinking) : No!

$$\text{i)} \quad \lambda = \begin{cases} \lambda_1, & t < \gamma \\ \lambda_2, & t \geq \gamma \end{cases} \quad \Rightarrow \gamma \text{ is a point change}$$

We believe there is a change point.

$$\text{ii)} \quad \lambda_j \sim \text{Exp}(\alpha)$$

Reasoning:  $\lambda \geq 0$  and continuous value

iii)  $\alpha$ : no further insights  $\rightarrow$  set to be a fixed number

$\gamma$ : could be any day  $\rightarrow \gamma \sim \text{DUnif}(1, 70)$

Posterior: • Where is change point?

$$P(\gamma | \xi, \alpha, \lambda)$$

↑ data

- What was the mean # of texts before and after?

$$P(\lambda | \gamma, \xi, \alpha)$$

Putting it all together:

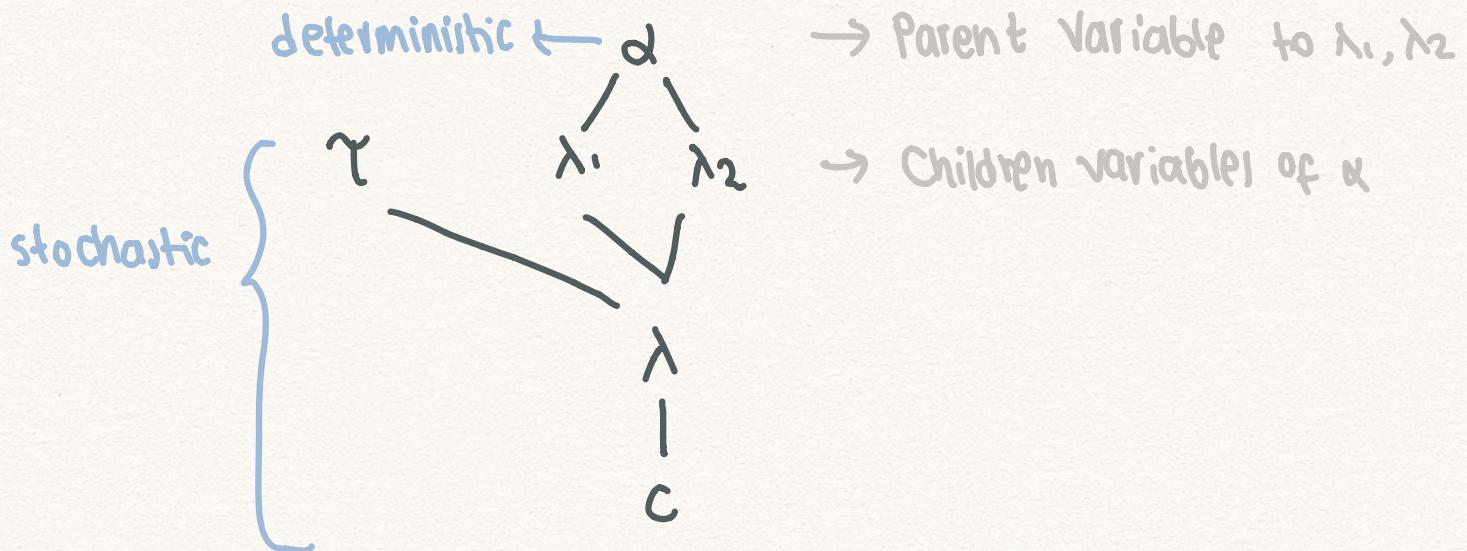
$$c_i \sim Po(\lambda)$$

$$\lambda = \begin{cases} \lambda_1, & t < \gamma \\ \lambda_2, & t \geq \gamma \end{cases}$$

$$\lambda_j \sim Exp(\alpha)$$

$$\gamma \sim Unif(0, 70)$$

\* If we want to know  $\gamma$ :



# 🥊 Stochastic vs Deterministic

- \* Stochastic variables are random: there are probabilities associated w/ each of its possible values.
- \* Deterministic variables are fixed numbers.

Example: Comparison

$$\textcircled{1} \quad \lambda \sim \text{Exp}(\alpha)$$

$$\textcircled{2} \quad \lambda = \alpha^2$$

Now we know  $\alpha = 0.5$

In case  $\textcircled{1}$ ,  $\lambda$  follows a distribution where we must sample (stochastic)

In  $\textcircled{2}$ ,  $\lambda = 0.5^2$  (deterministic)

- \* The `.random()` function: Sample from the distribution specified.

Say we have posterior values for  $\lambda_1$  and  $\lambda_2$  (call them `lambda-1` and `lambda-2`)

I want to know what is  $|\lambda_2 - \lambda_1|$

↳ what is the magnitude of the average change.

Ans: Draw one sample from each distribution, take their difference, then the absolute value. Do it many times!

(Key point for exam! :p)

In Pseudo code:

```

for i=1 : 10000
    diff[i] = abs(lambda_2.random(i) - lambda_1.random(i))
end

```

posterior for  $|\lambda_2 - \lambda_1|$  is diff.

(Powerful! Mathematically would be hard to compute).

\* So with posteriors of variables  $\lambda_1$  and  $\lambda_2$ , we can easily obtain a posterior for any function  $f(\lambda_1, \lambda_2)$  by sampling.

↳ simple case

(?) Julia, independence

Post. for  $\lambda_1$ :  $P(\lambda_1 | \Sigma, \alpha, \lambda_2, \gamma)$

Post. for  $\lambda_2$ :  $P(\lambda_2 | \Sigma, \alpha, \lambda_1, \gamma)$

Hai: If you know the dependence, you only need to sample from one.

Quick view:

$$c_i \sim Po(\lambda)$$

↓

$$\lambda = \begin{cases} \lambda_1 \\ \lambda_2 \end{cases}$$

$$\lambda_j \sim Exp(\alpha)$$

Specify  $\alpha$   
(maybe using data)

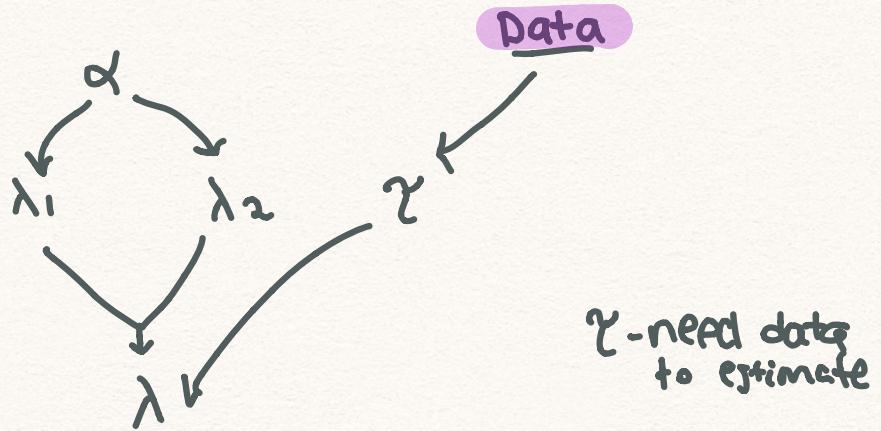
↓

$$\lambda_j \sim Exp(\alpha)$$

↳ where is  $\gamma$ ?

↓  
We need data to answer this question.

Nice view:



Why is  $\gamma$  stochastic?

$$\gamma \sim \text{Unif}(1, 20)$$

$$P(\gamma | \lambda_1, \lambda_2, \alpha, \xi)$$

•  $\gamma$  doesn't depend on  $\lambda_1, \lambda_2$ . (It's the other way around)

$$\Rightarrow P(\gamma | \alpha, \xi) \Rightarrow P(\gamma | \xi)$$

$$\hookrightarrow P(\lambda_1, \lambda_2 | \gamma, \alpha, \xi)$$

- \* For simulation, each variable only needs its parents variables.
- \* For estimation, we also need data.

## 2) BAYESIAN MODELING + APPLICATIONS

**BIG QUESTION:** What underlying process gave rise to our observed data?

1) What model is a good model for the observed data?

(This is data generating distribution)

2) Does the model in 1 have parameters?

(Probably if using a stat model)

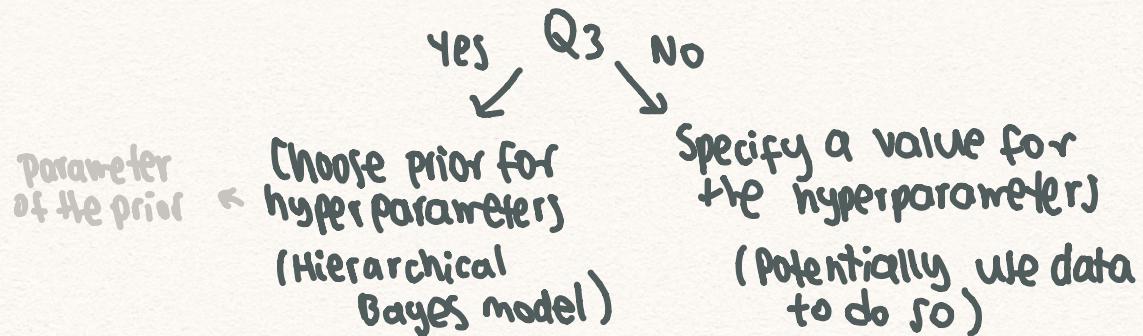
↳ If so, do we know anything about how those parameters were generated?

(Prior model) If no, follow frequentists point of view.

3) Does the model in ② have Parameters?

(Probably)

↳ Q3: Do we know anything about them?



Next time: Example) Challenger + truth serum  
Goodness of fit  
Talk about the exam, solve question