Date: ~~Tuesday~~ Thursday, 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: Poi(λ), Exp(α), U(a,b), DiscUnif(a,b), Normal(μ, σ^2), Bin(n,p)

    ↳ what are they used for? so on...

  2. Have class T, Th, F

     Friday times: 10:00-11:55, 2:00-3:55
- HW due tonight

3 people have the same birthday in session 1!

Aug 28: Mawa, Tianqi, 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:
- $Bin(n, p)$
- $Normal(\mu, \sigma^2)$

1) Working with PyMC

Example: Text messages

Data: C_i → count of incoming messages on day i

$C_i \sim Po(\lambda)$

↳ mean count (but also the variance, specific to this distribution).

Prior: Do we know anything about λ?

(Bayesian way of thinking): Yes!

(Frequentist way of thinking): No!

i) $\lambda = \begin{cases} \lambda_1, & t < \tau \\ \lambda_2, & t \geq \tau \end{cases}$ → τ is a point change

we believe there is a change point.

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

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

iii) α: no further insights → set to be a fixed number

τ: could be any day → $\tau \sim DUnif(1, 70)$

Posterior:

- Where is change point?

  $\mathbb{P}(\tau \mid \underset{\sim}{c}, \alpha, \lambda)$

  ↳ data

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

  $\mathbb{P}(\lambda \mid \tau, \underset{\sim}{c}, \alpha)$

Putting it all together:

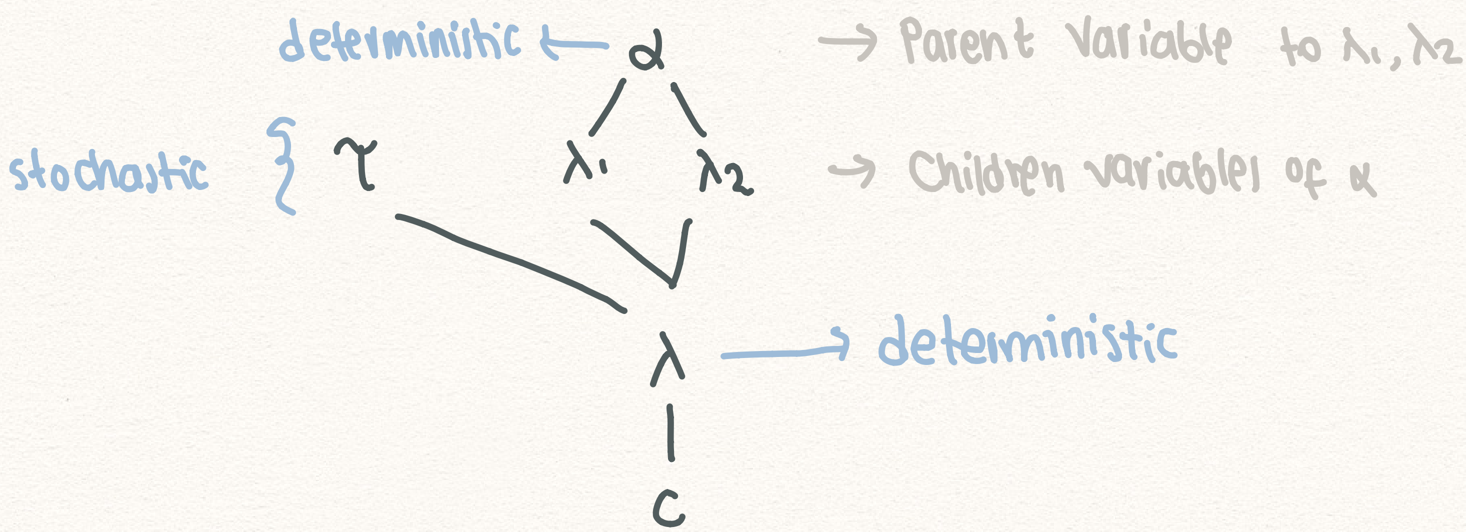
$$C_i \sim Po(\lambda)$$

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

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

$$\tau \sim Dunif(0, 70)$$

* If we want to know τ:



🥊 Stochastic vs Deterministic

* Stochastic variables are random: there are probabilities associated w/ each of its possible values.

* Deterministic variables are fixed numbers.

Example: Comparison

① $\lambda \sim Exp(\alpha)$

② $\lambda = \alpha^2$

Now we know $\alpha = 0.5$

In case ①, λ follows a distribution where we must sample (stochastic)

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

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

Say we have posterior values for λ_1 and λ_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(1) - lambda_1.random(1))
end
```

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

(Powerful! Mathematically would be hard to compute).

* So with posteriors of variables λ_1 and λ_2, we can easily obtain a posterior for any function $f(\lambda_1, \lambda_2)$ by sampling.

↳ simple case
(?) Julia, independence

post. for λ_1: $\mathbb{P}(\lambda_1 | \underline{C}, \alpha, \lambda_2, \tau)$

post. for λ_2: $\mathbb{P}(\lambda_2 | \underline{C}, \alpha, \lambda_1, \tau)$

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 α
(maybe using data)

↓

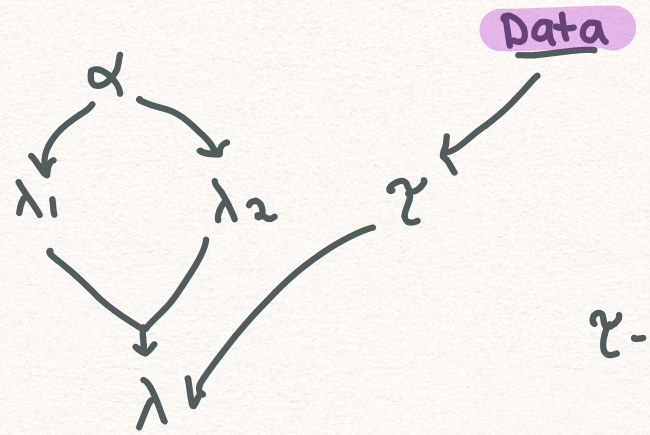
$\lambda_j \sim Exp(\alpha)$

↳ where is τ?

↓

We need data to answer this question.

Nice view:



τ - need data to estimate

Why is τ stochastic?

$\tau \sim \text{DUnif}(1, 20)$

$P(\tau \mid \lambda_1, \lambda_2, \alpha, \underset{\sim}{c})$

- τ doesn't depend on λ_1, λ_2. (It's the other way around)

$\Rightarrow P(\tau \mid \alpha, \underset{\sim}{c}) \Rightarrow P(\tau \mid \underset{\sim}{c})$

$\hookrightarrow P(\lambda_1, \lambda_2 \mid \tau, \alpha, \underset{\sim}{c})$

* 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?

Q3

Yes

Choose prior for hyperparameters

(Hierarchical Bayes model)

Parameter of the prior

No

Specify a value for the hyperparameters

(Potentially use data to do so)

Next time: Examples Challenger + truth serum
Goodness of fit
Talk about the exam, solve questions