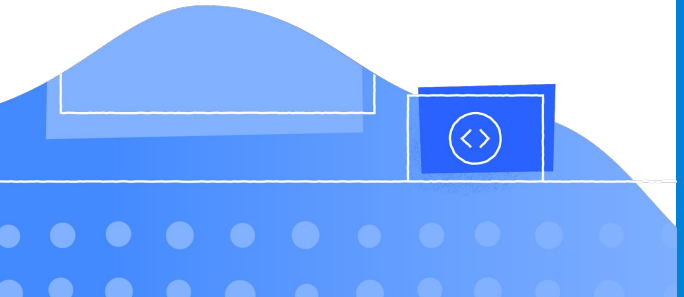
GDG

YYC DATA SOCIETY

Ben Reeves
Viewpoint Investment Partners
YYC Data Society

Tell your Data's Story; the Bayesian Way!

human intelligence + machine intelligence = ai nirvana



What to do when your data wants to lie to you?

Low Signal to Noise

+

Few/Inconsistent Examples

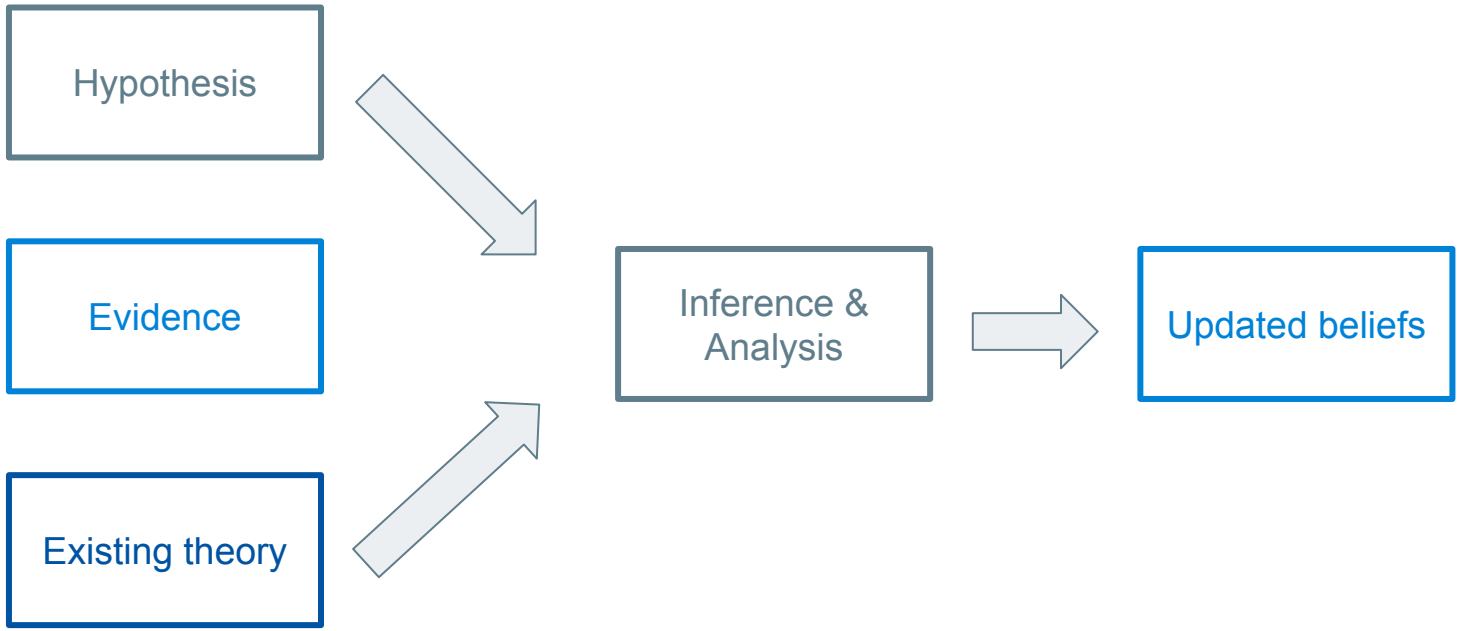
+

Data changing over time

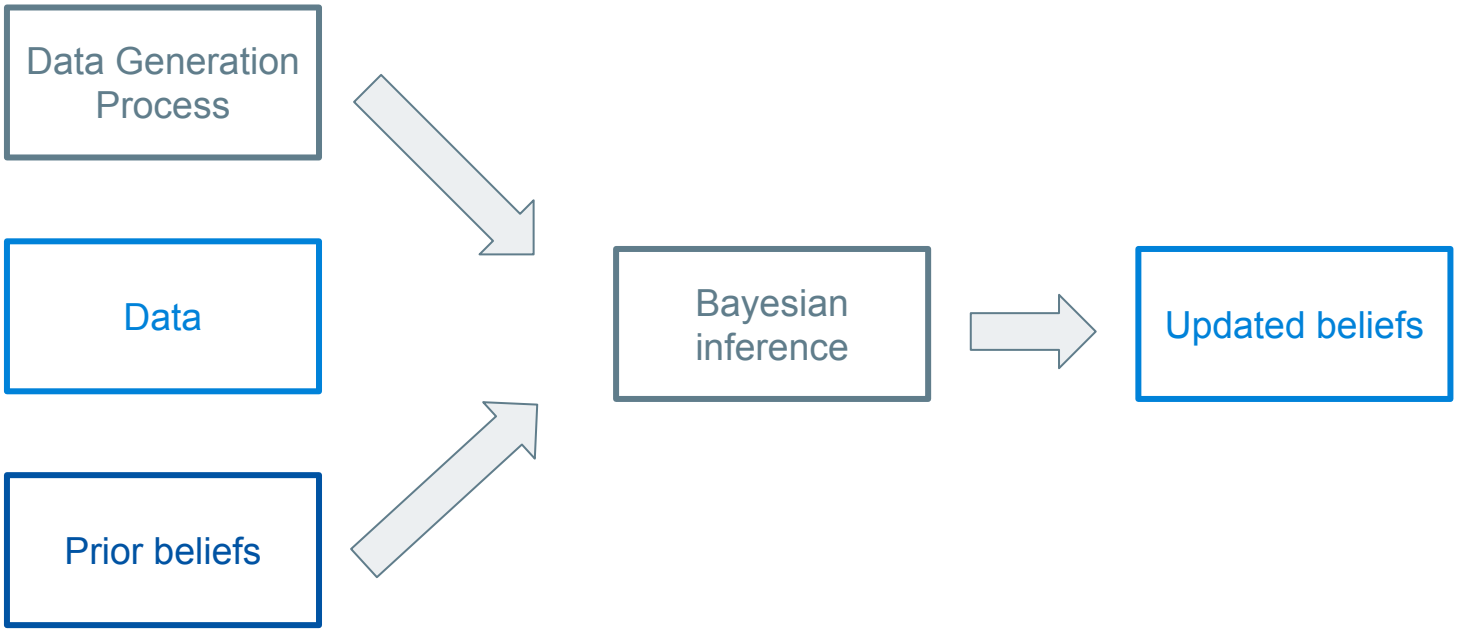
=

Unkind Learning Environment

The scientific method: one tool to fight the chaos

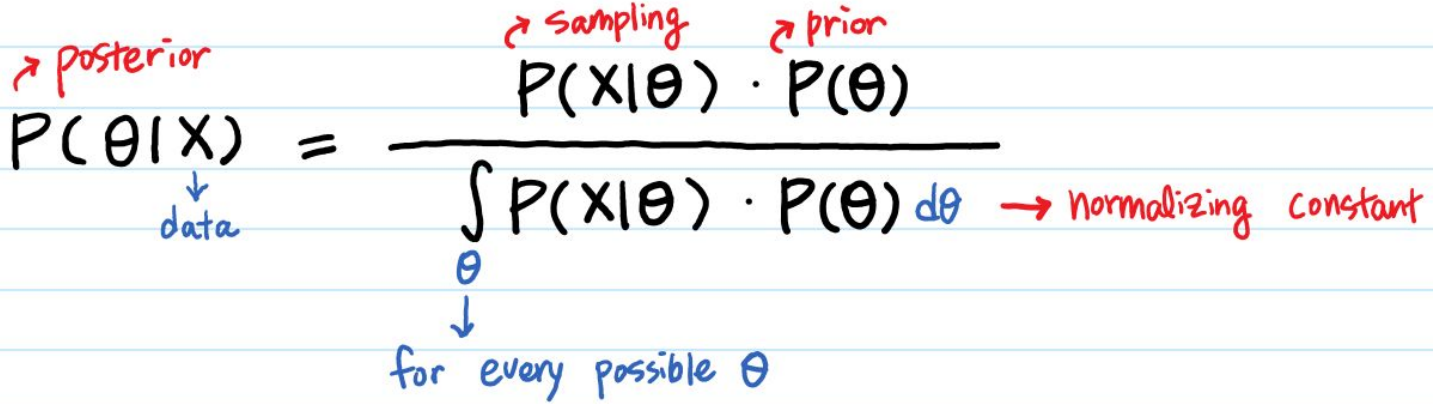


Putting the science in data science



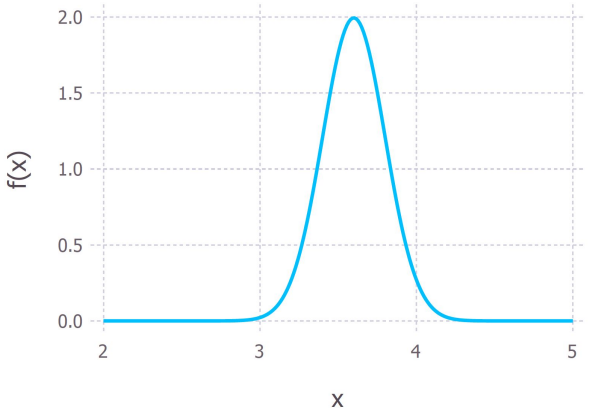
A Motivating Example

Breaking Down Bayes

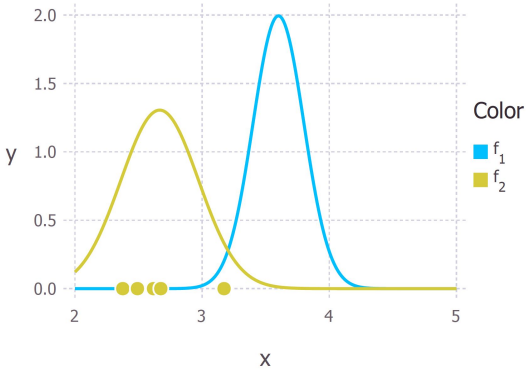


Bayes: just a weighted average of hypothesis and evidence

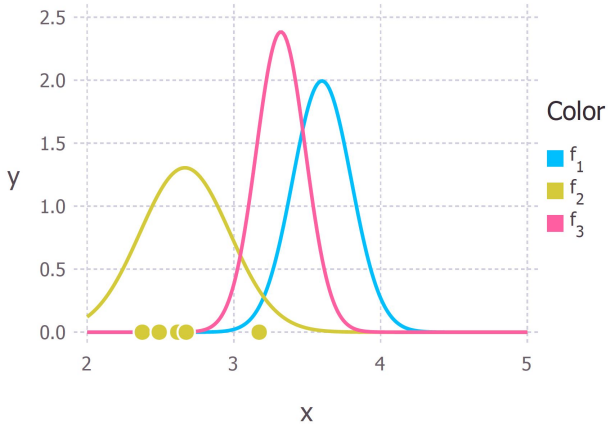
Prior beliefs [Blue]
I think 3%-4% of the population will get Covid vaccines per month



Data [Yellow]
We get vaccinated, but not as much as expected

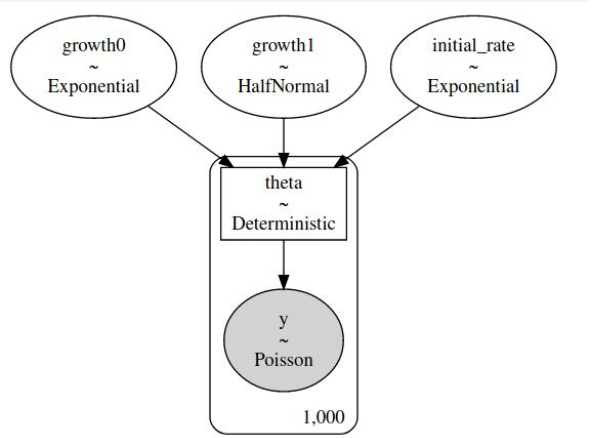


Posterior [Pink]
I still think we'll get vaccinated faster, but my expectations are tempered



Back to our CMO's question

- You have intuition about what might have happened
- We will turn that intuition into a story and test its veracity

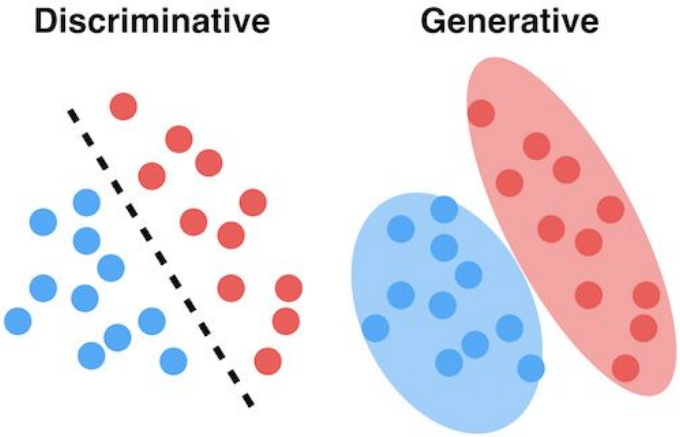


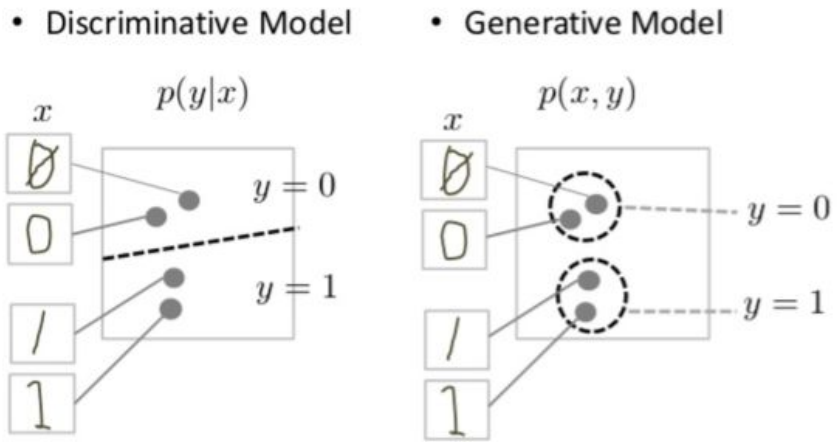
MCMC Demo

https://colab.research.google.com/drive/1kIVWy9rBRCMlZvTzwGupkaqLpFD45rl8

A Tale of Two Models

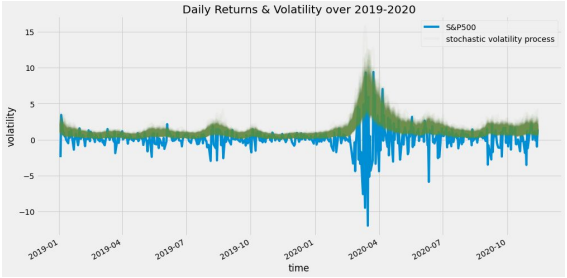
Drawing boundaries vs modelling distributions

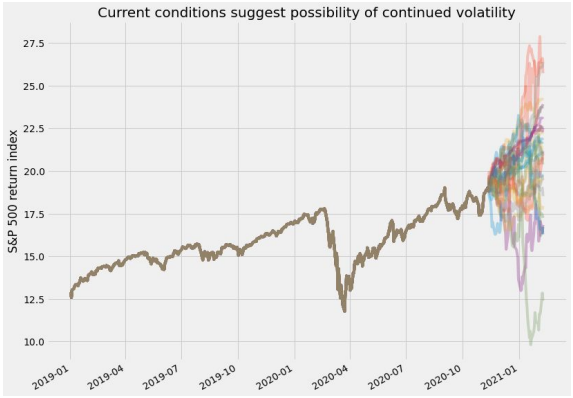




Understanding the past to simulate the future

- Fit a data generation process to S&P 500 returns
- Check reasonability of our model
- Simulate future possibilities given present conditions





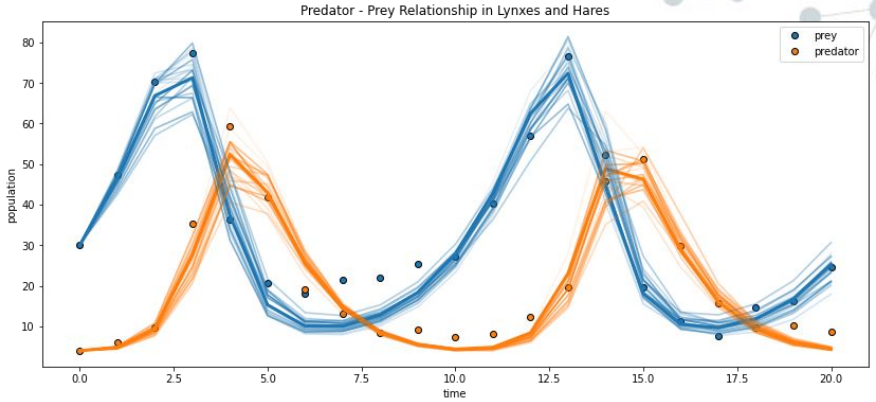
S&P500 Demo

https://colab.research.google.com/drive/1w0HPapmvPLNIngi1801BmdpbqFYQn8XL#scrollTo=ivmdm53UEs4Y

YYC DATA

Testing scientific hypotheses with Bayesian statistics

- Predator-prey relationship expressed as ODE
- Wanted: estimate parameters from data
- Simulate possibilities and search for best parameters



$$\frac{\mathrm{d}}{\mathrm{d}t}u = (\alpha - \beta v)u \quad = \alpha u - \beta uv$$

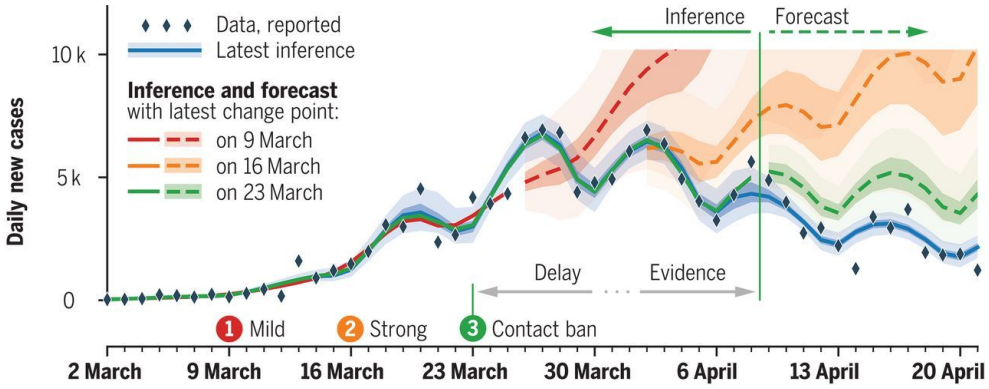
$$\frac{\mathrm{d}}{\mathrm{d}t}v = (-\gamma + \delta\, u)\, v \quad = -\gamma v + \delta uv$$

ODE Demo!

https://colab.research.google.com/drive/1CqlZDVHRC7Yqfkw6EYHhGWELJ3og0GtI

Conclusion: All models are approximations

- Useful models answer questions:
  - What if I see X?
  - What if Y changes?
  - How likely is Z?
- Generative models help us understand how our data came to be
- Bayesian methods help our models understand how our data came to be



Thanks!

Shameless Plugs:

- YYC DataCon
- YYC Data Society