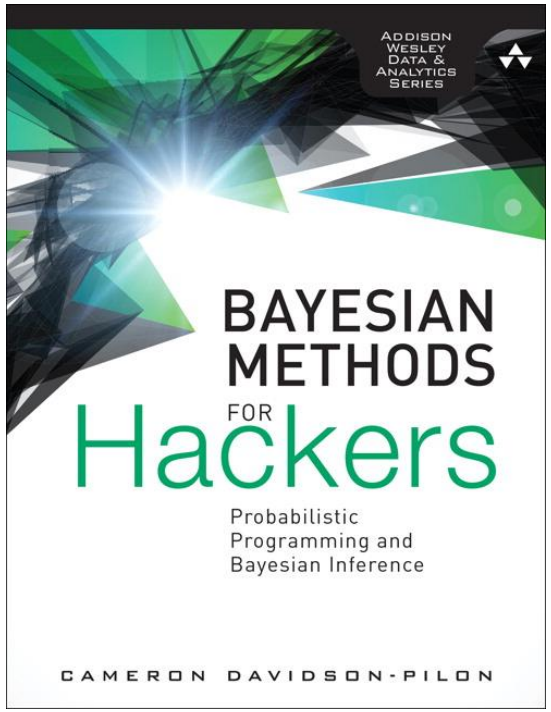
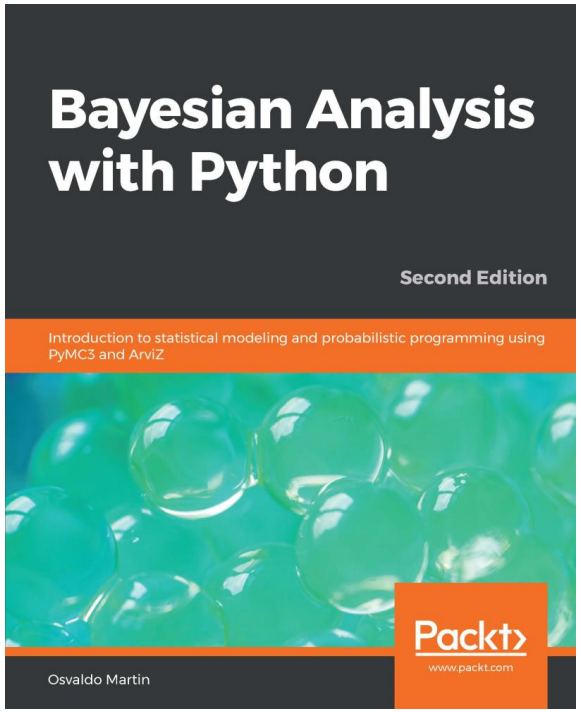
Bayesian modeling

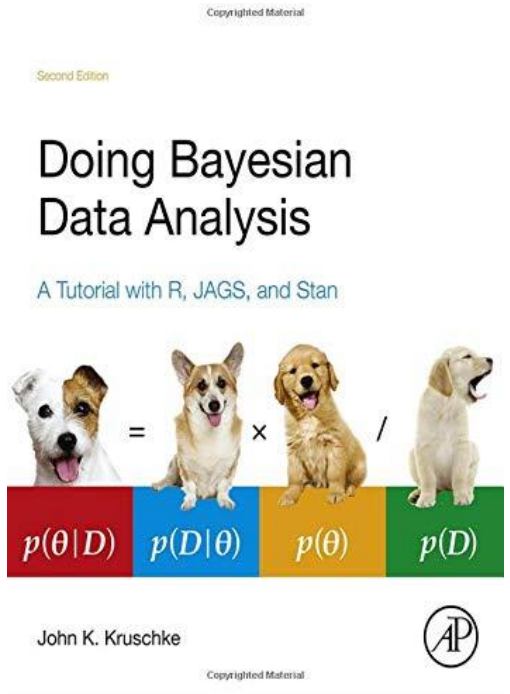
Thinking probalistically

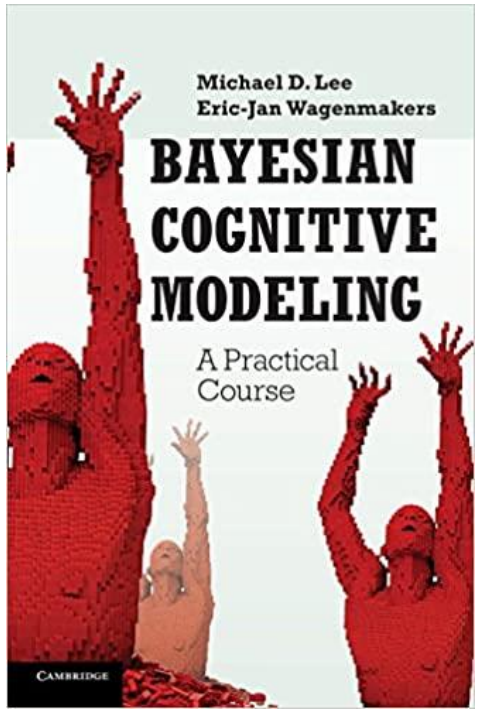
Introduction to PyMC3

Thinking probabilistically









Thinking probabilistically

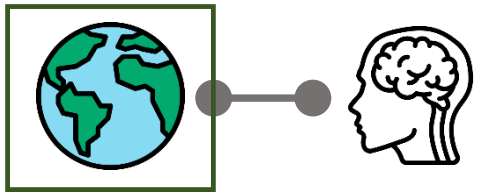
Frequentist

Bayesian

Probability is the long-run frequency of events.

$$f(y|\theta)$$

Fixed

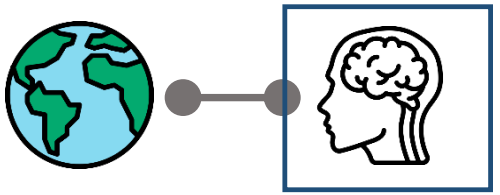


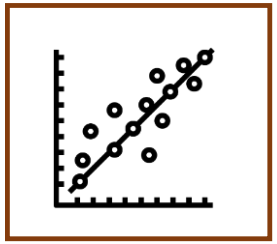
Probability measure the *believability in an event*.

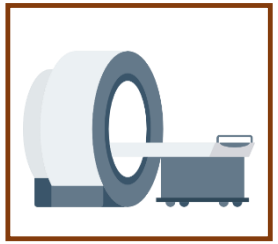
$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

Fixed

Bayesian inference is simply updating your beliefs after considering new evidence.

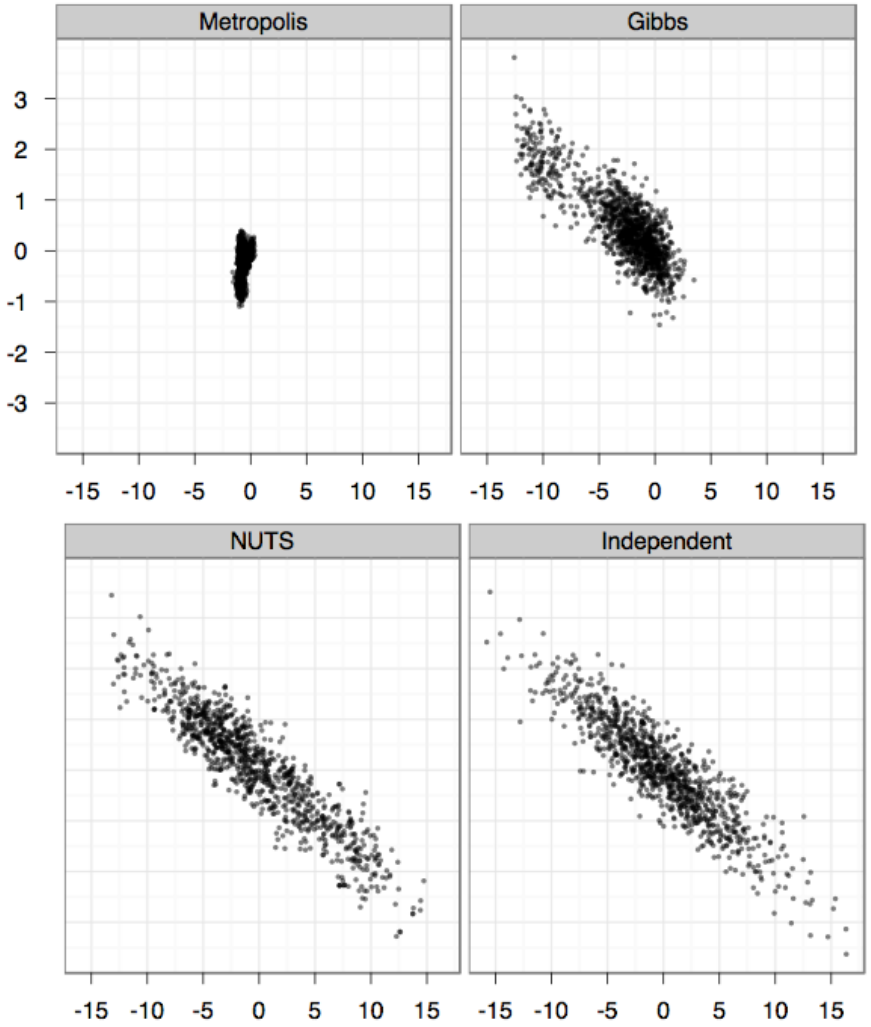






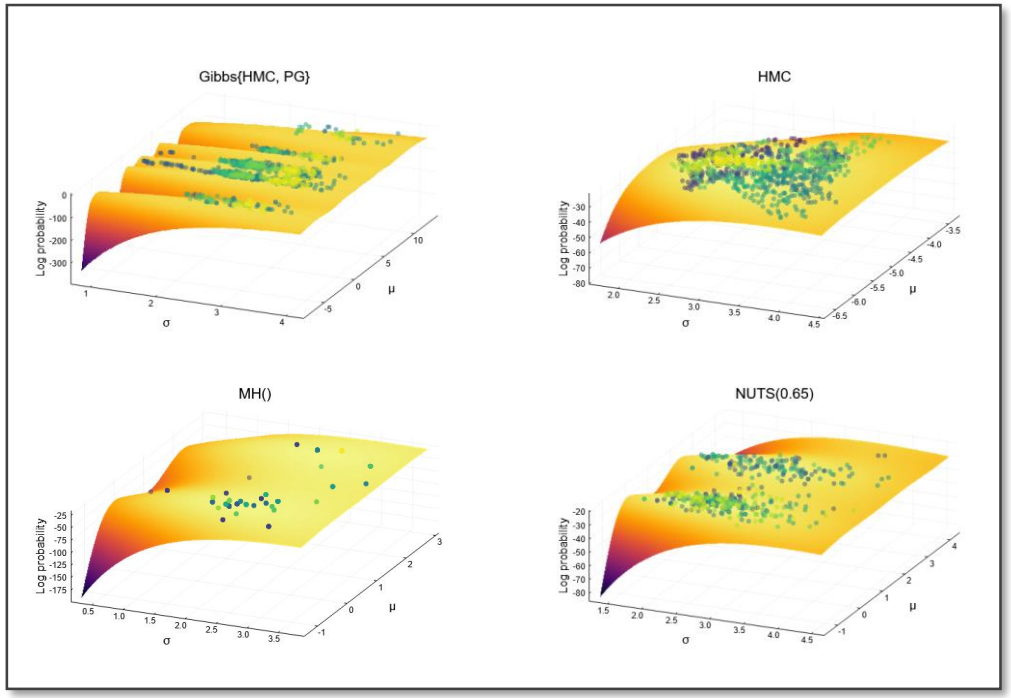
MCMC samplers

- Approximation
  - Variational inference
- Stochastic sampling
  - MCMC methods



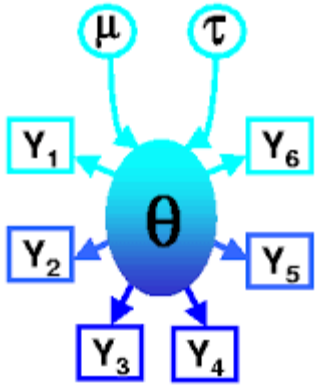
Online demo:

https://chi-feng.github.io/mcmc-demo/



https://turing.ml/dev/docs/using-turing/sampler-viz

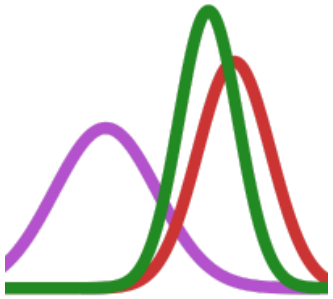
Probabilistic programming languages



BUGS

JAGS

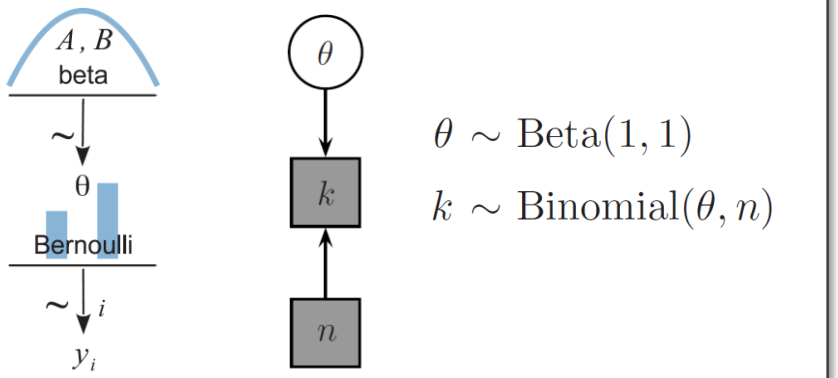
Stan



Turing.jl

Probabilistic programming languages

Graphical model

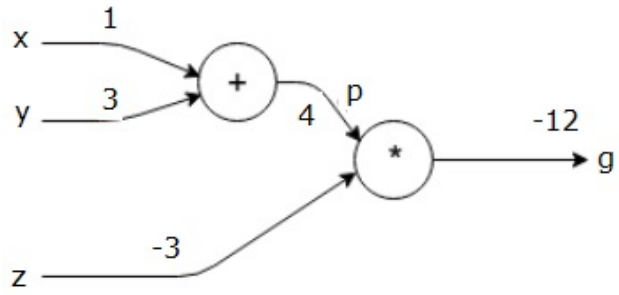


$$\theta \sim \text{Beta}(1, 1)$$

$$k \sim \text{Binomial}(\theta, n)$$

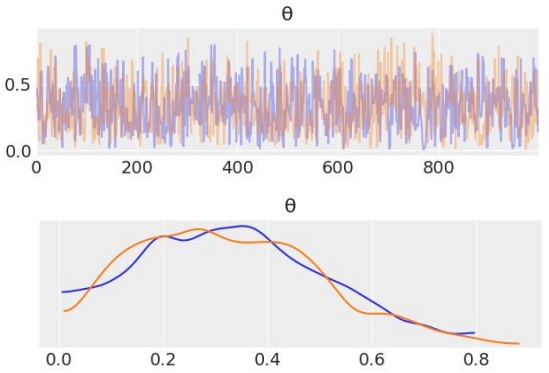
```
with pm.Model() as our_first_model:
  θ = pm.Beta('θ', alpha=1., beta=1.)
  y = pm.Bernoulli('y', p=θ, observed=data)
  trace = pm.sample(1000, random_seed=123)
```

Computational graph



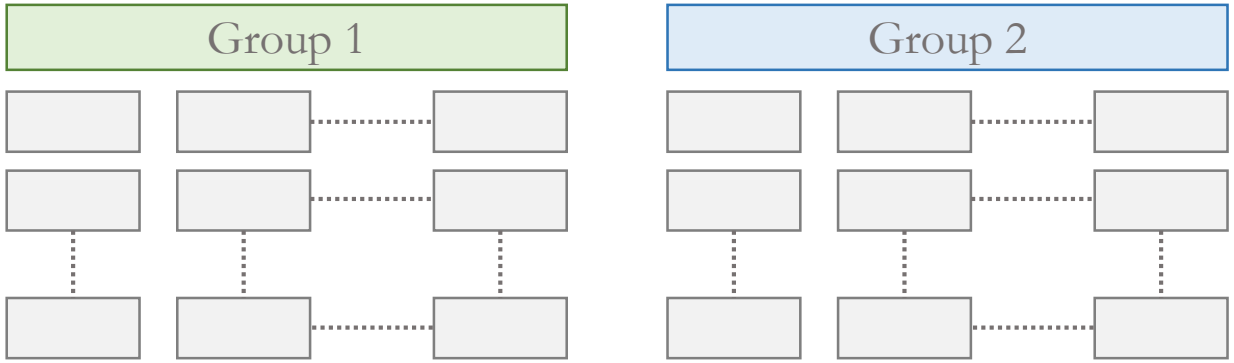
theano

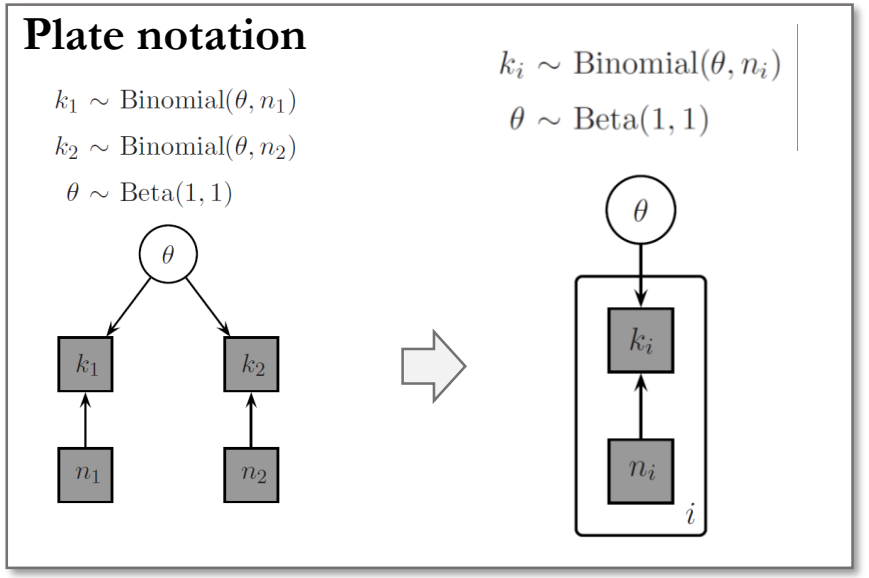
- Automatic differentiation
- GPU computing
- Optimizations



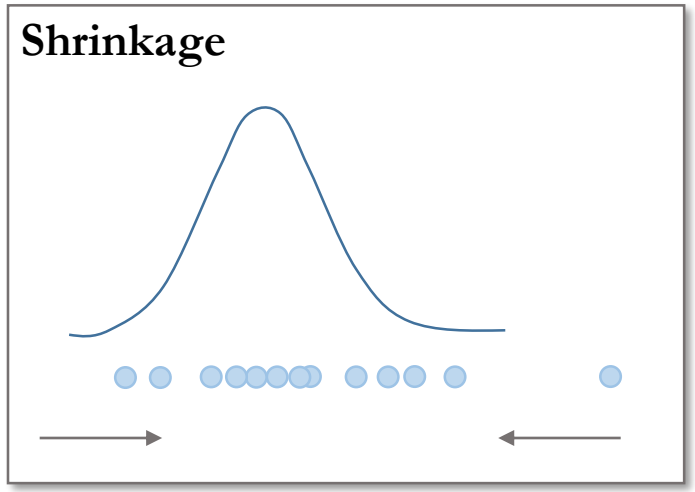
Sampling

Hierarchical/multilevel models



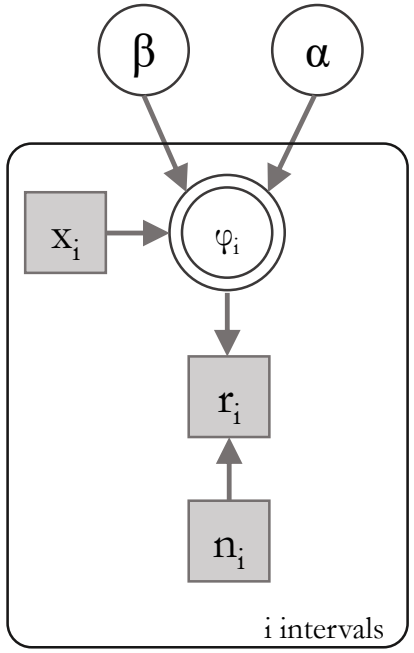


- Hyperpriors
- Hyperparameters

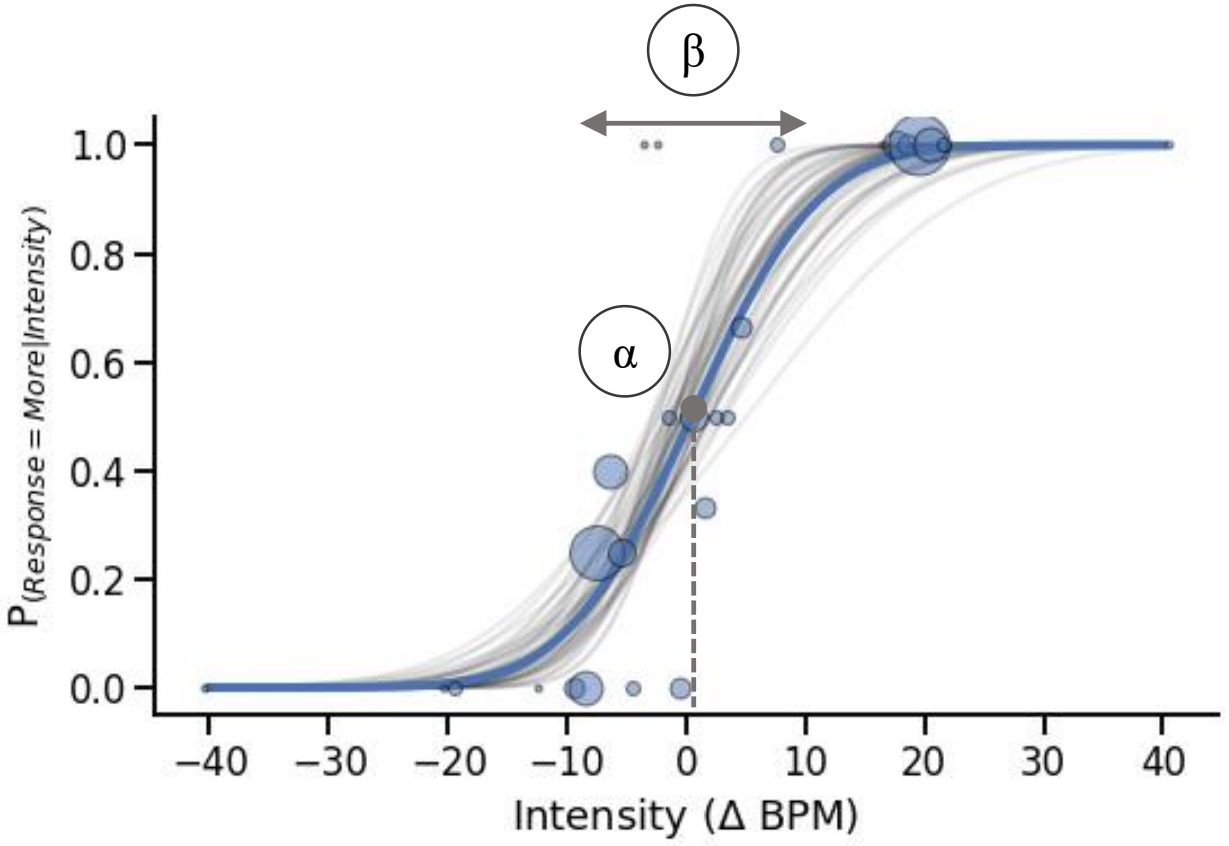


Psychophysics

Psychophysics is concerned with measuring how external physical stimuli cause internal psychological sensations.

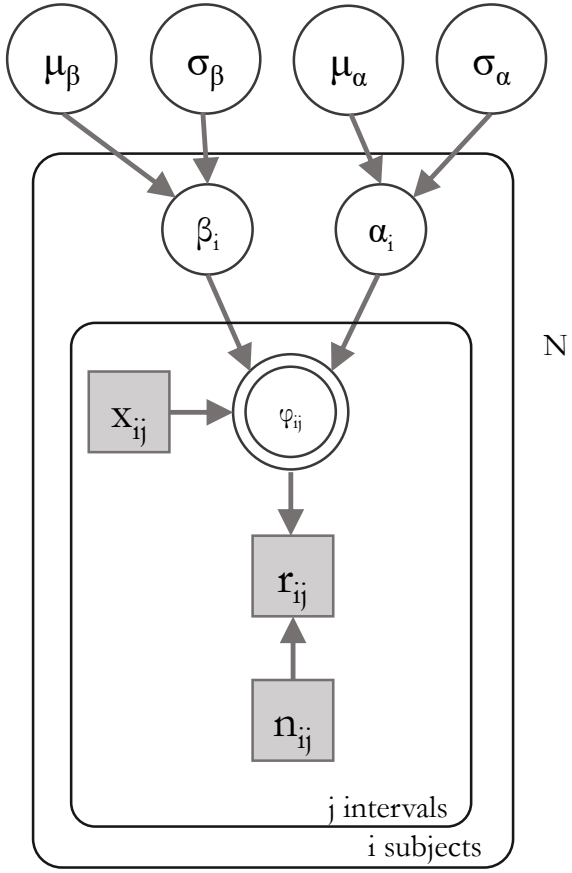


N responses	r_i	=	0	3	10 .....
N trials	n_i	=	5	20	15 .....
Intensities	x_i	=	-20	-10	0 .....

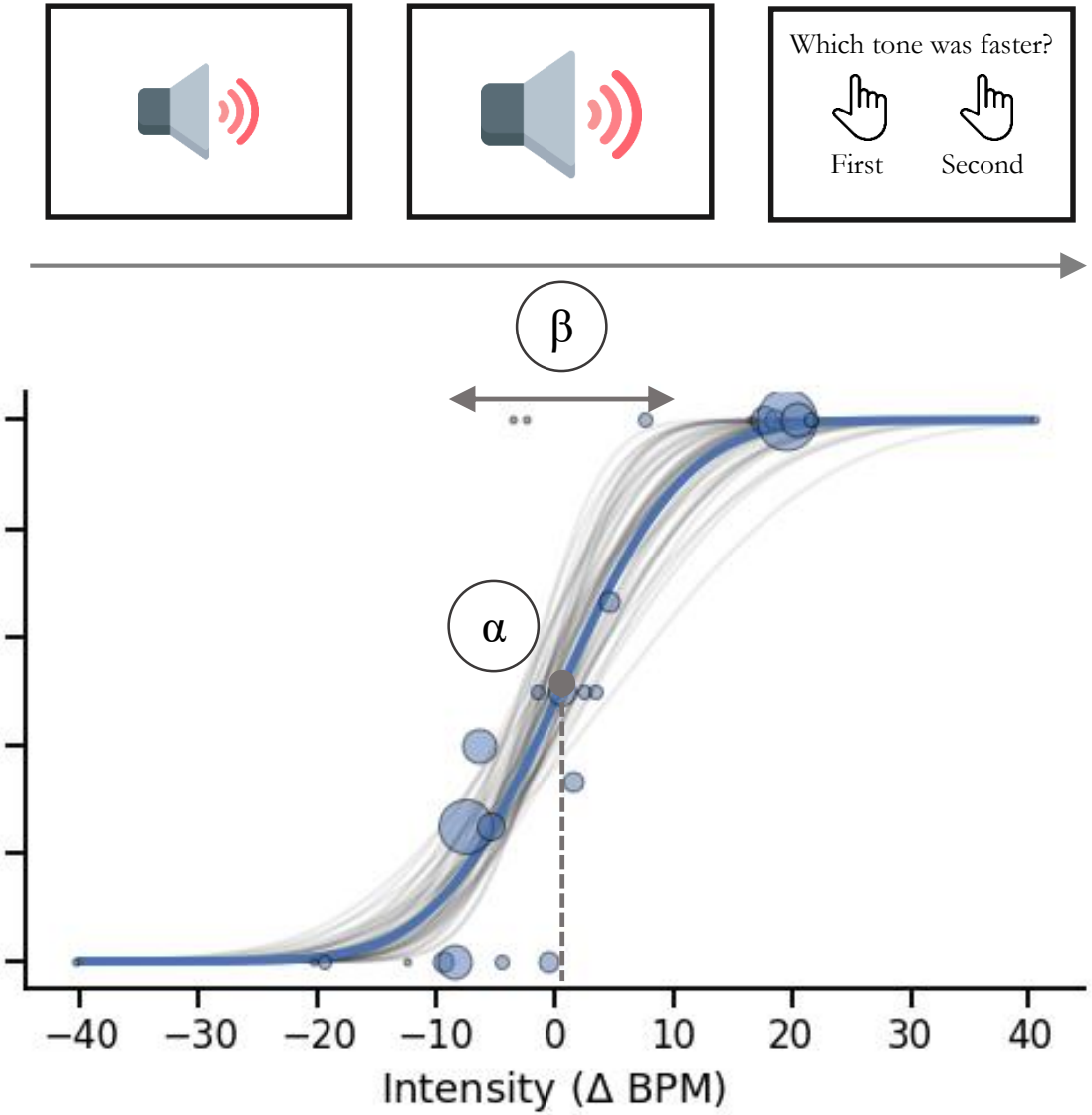


Hierarchial psychophysics

Psychophysics is concerned with measuring how external physical stimuli cause internal psychological sensations.

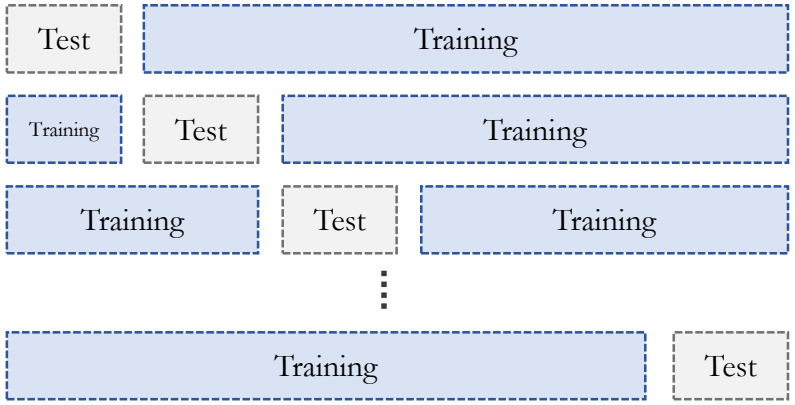


N responses	r_i	=	0	3	10 .....
N trials	n_i	=	5	20	15 .....
Intensities	x_i	=	-20	-10	0 .....
Subject ID	S_i	=	0	0	0 .....



Model comparison

Cross-validation



Information criteria

How well models fit the data while taking into account their complexity through a penalization term

- Mean square error (MSE)

$$\frac{1}{n}\sum_{i=1}^{n}(yi - E(y_i \mid \theta))^2$$

- Log-likelihood

$$\sum_{i=1}^{n}\log p(y_i \mid \theta)$$

- Deviance

$$-2\sum_{i=1}^{n}\log p(y_i \mid \theta)$$

- Akaike information criterion

$$AIC = -2\sum_{i=1}^{n}\log p(y_i \mid \theta_{mle}) + 2pAIC$$

How well the model fits the data — Penalizes complex models

If we have two or more equi valent explanations for the s ame phenomenon, we shoul d choose the simpler one... but also the more accurate.

- Widely applicable information Criterion

$$WAIC = -2\text{lppd} + 2p_{WAIC}$$

How well the model fits the data — Penalizes complex models

• Nicolas Legrand • Postdoctoral fellow • Embodied Computation Group

Resources

Chris Fonnesbeck - An introduction to Markov Chain Monte Carlo using PyMC3 | PyData London 2019
https://www.youtube.com/watch?v=SS_pqgFziAg

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. Nature Reviews Methods Primers, 1(1). https://doi.org/10.1038/s43586-020-00001-2

Kruschke, J. (2015). Doing Bayesian data analysis : a tutorial with R, JAGS, and Stan. Boston: Academic Press.