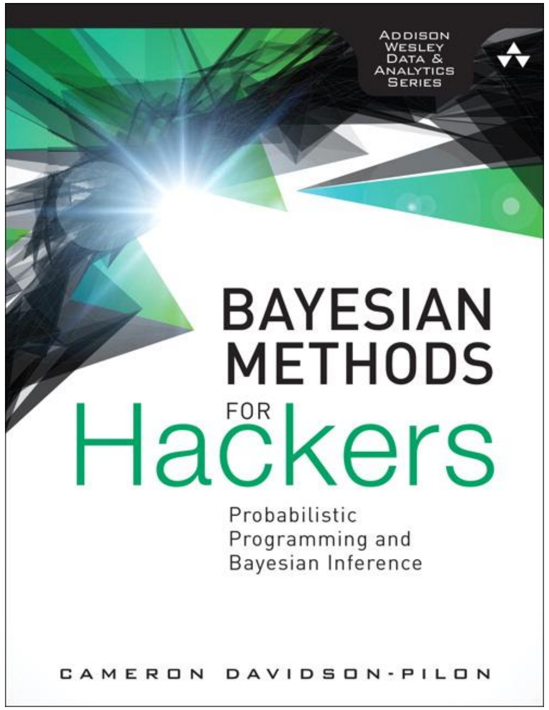
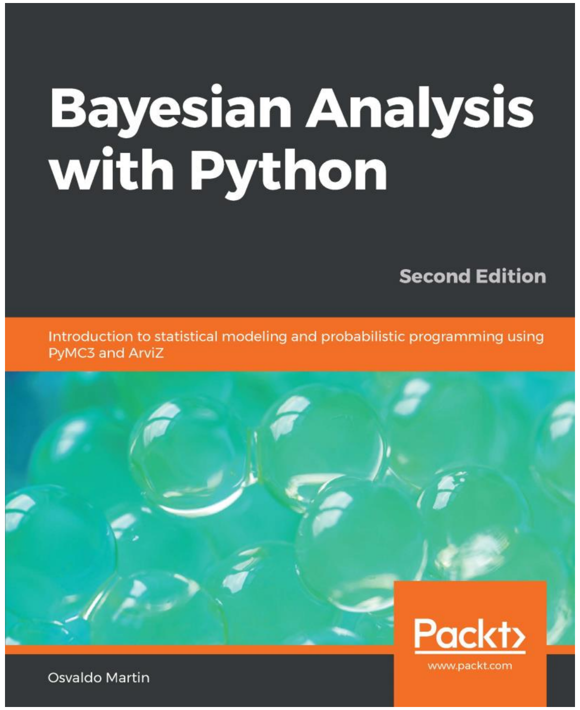
Bayesian modeling

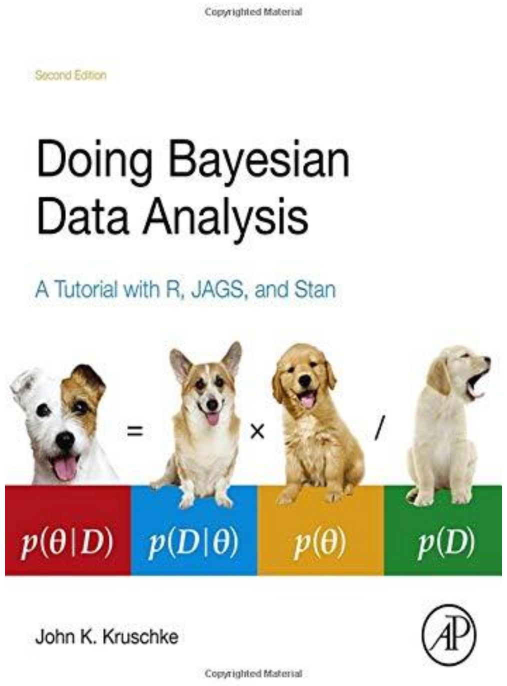
Thinking probalistically

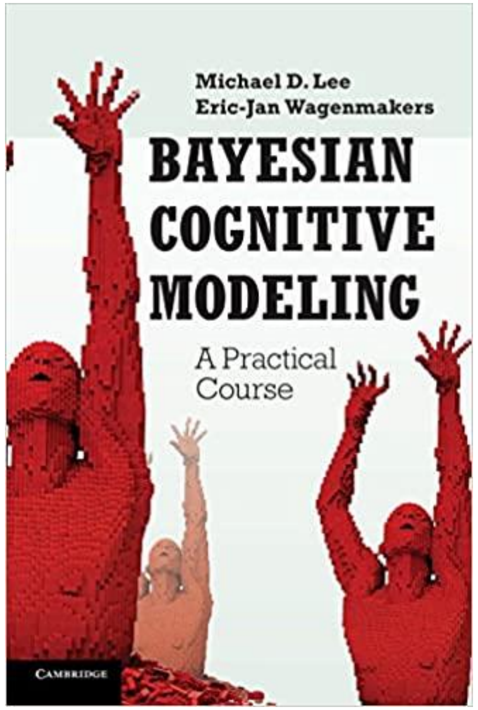
Introduction to PyMC3

Thinking probabilistically









Thinking probabilistically

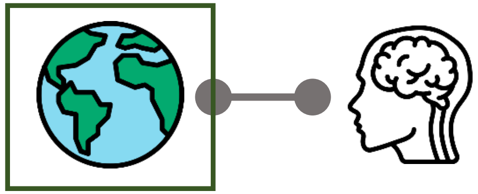
Frequentist

Bayesian

Probability is the long-run frequency of events.

$f(y|\theta)$

Fixed

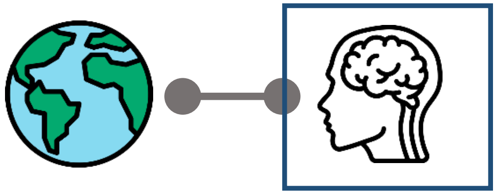


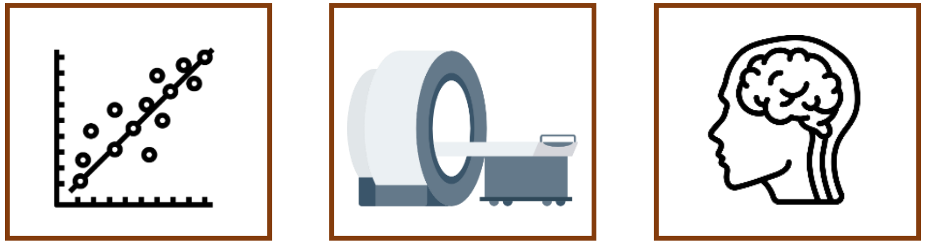
Probability measure the *believability in an event*.

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

Fixed

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





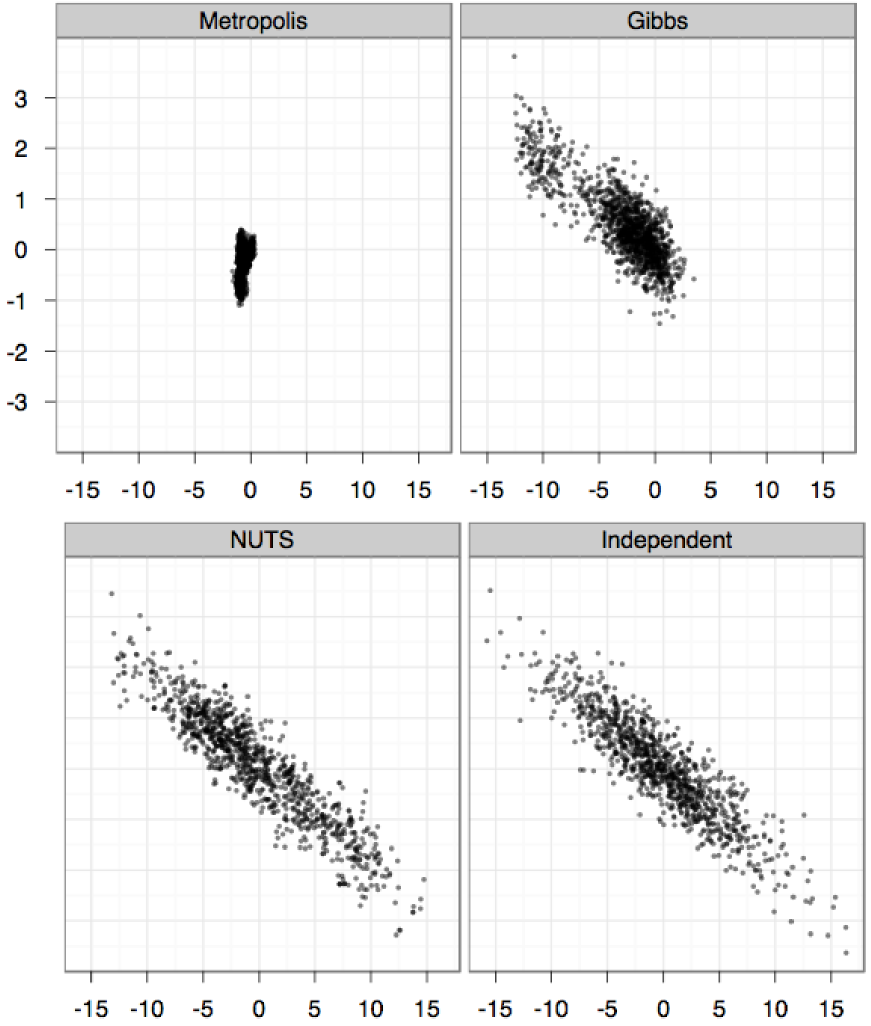
• Nicolas Legrand • Postdoctoral fellow • Embodied Computation Group

ECG embodied computation group

AARHUS UNIVERSITY

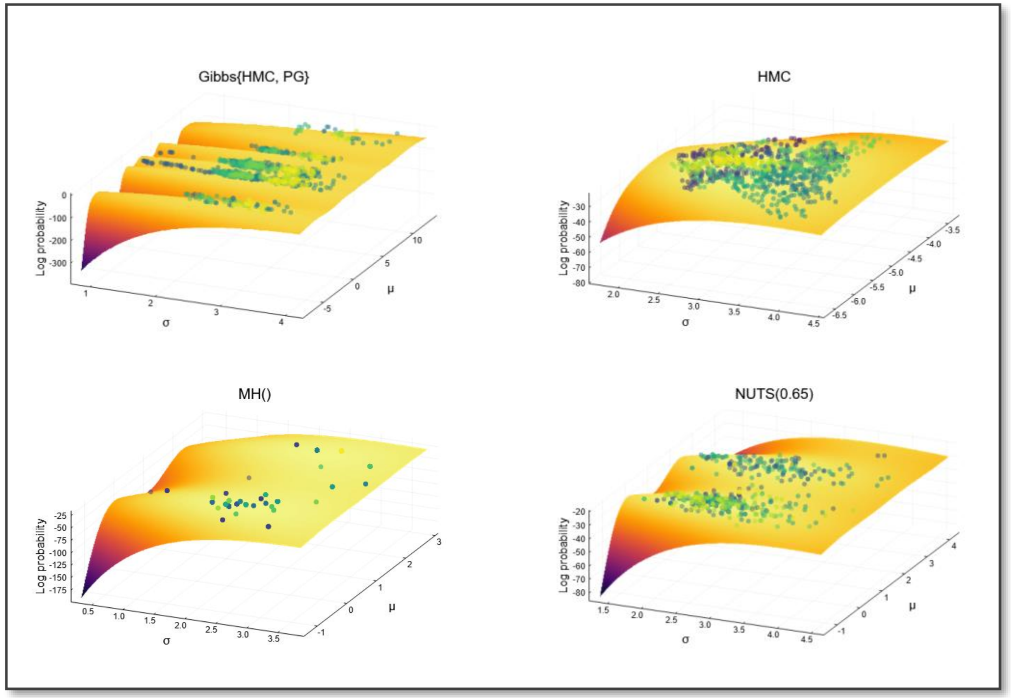
MCMC samplers

- 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

BAYESIAN MODELING WITH RJAGS

PyMC3

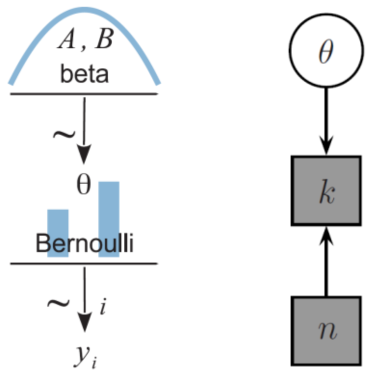
Stan

Turing.jl

TensorFlow

Probabilistic programming languages

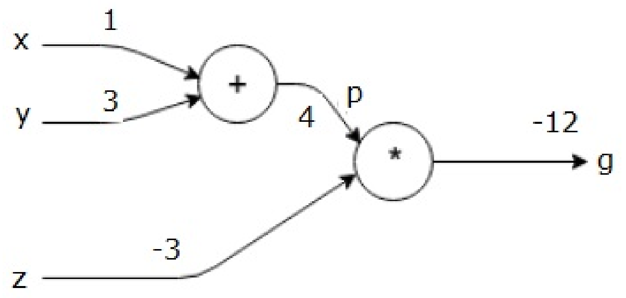
Graphical model



$$\theta \sim \mathrm{Beta}(1,1)$$
$$k \sim \mathrm{Binomial}(\theta, n)$$

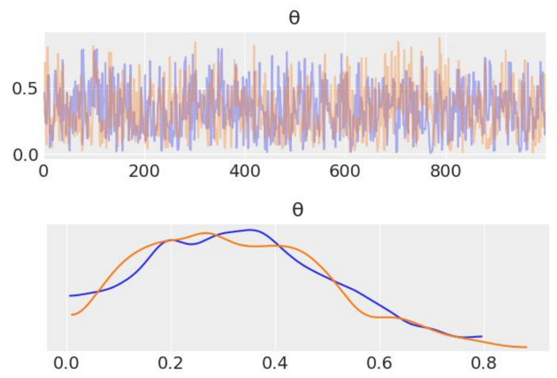
```python
with pm.Model() as our_first_model:
  θ = pm.Beta('θ', alpha=1., beta=1.)
  y = pm.Bernoulli('y', p=θ, observed=data)
  trace = pm.sample(1000, random_seed=123)
```

Computational graph



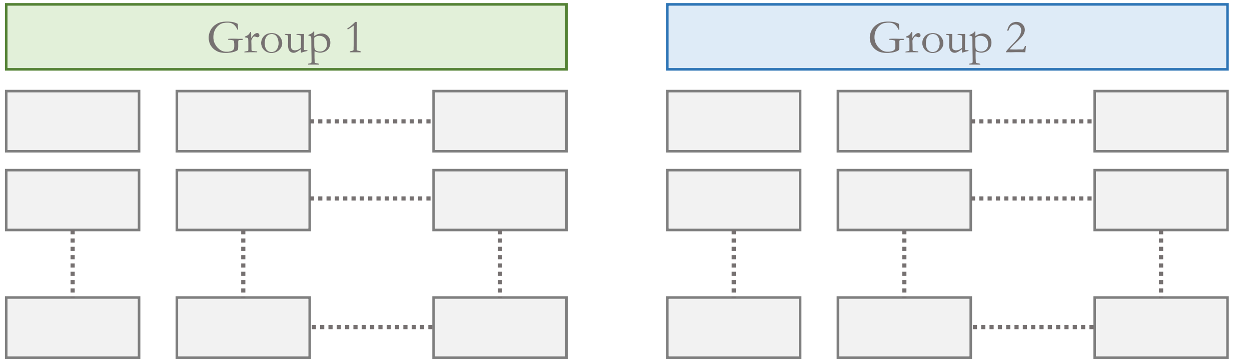
theano

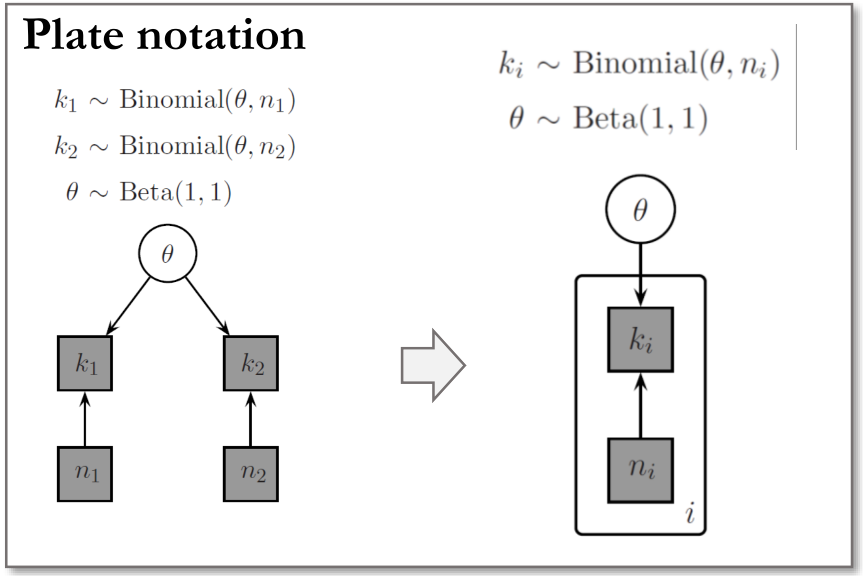
- Automatic differentiation
- GPU computing
- Optimizations



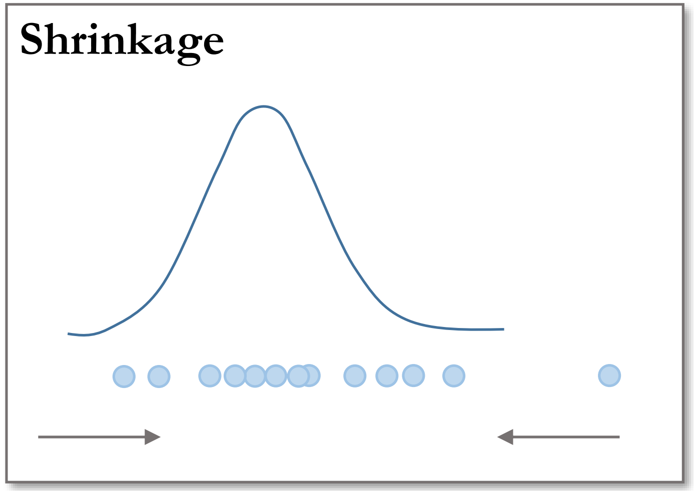
Sampling

Hierarchical/multilevel models



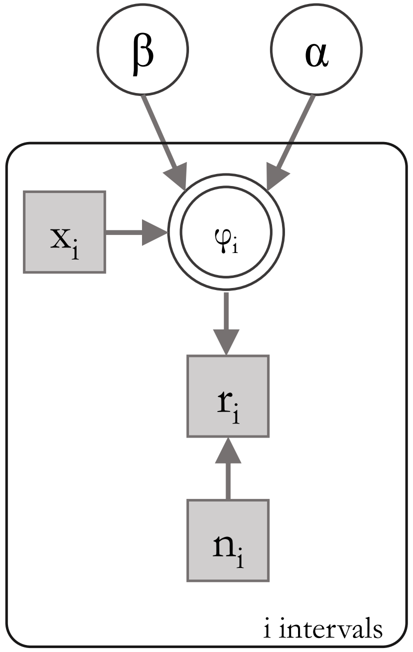


- Hyperpriors
- Hyperparameters



Psychophysics

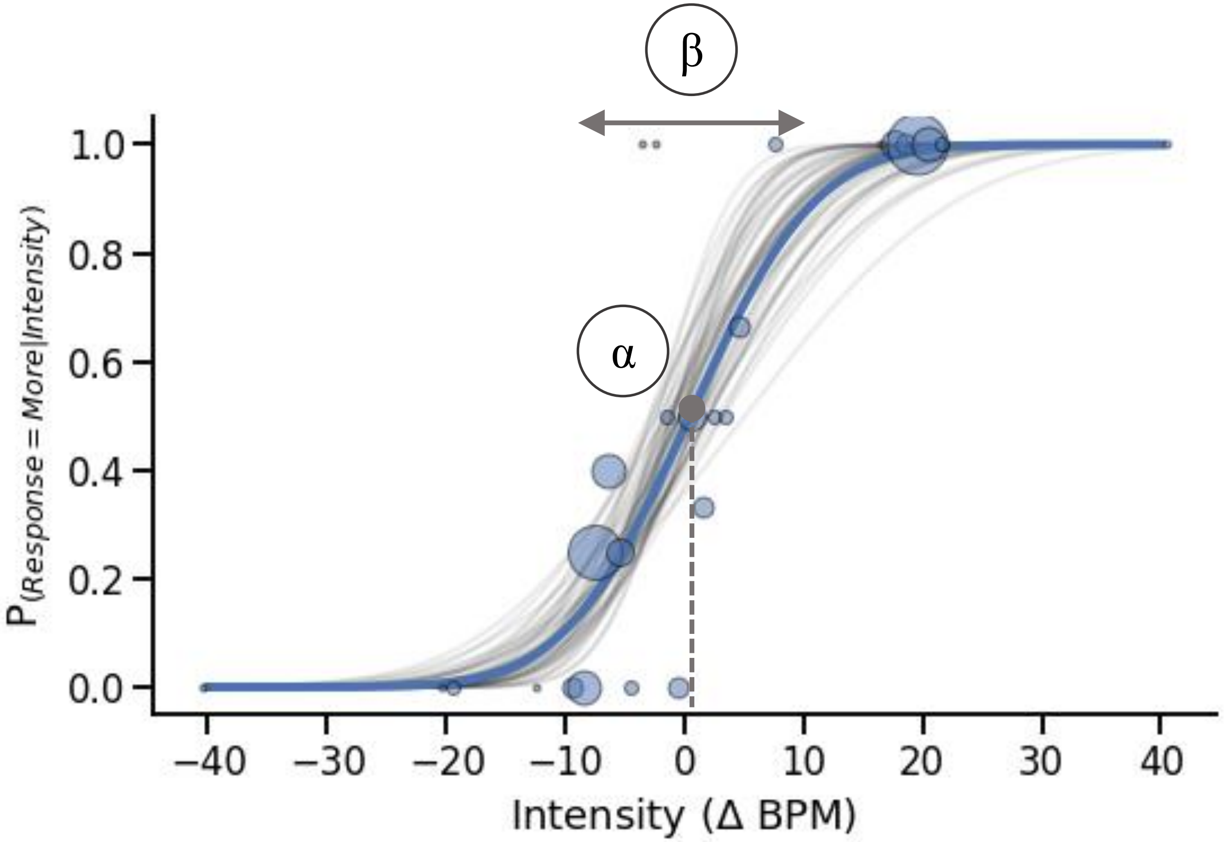
Psychophysics is concerned with measuring how external physical stimuli cause internal psychological sensations.



N responses r_i = 0 3 10 ....

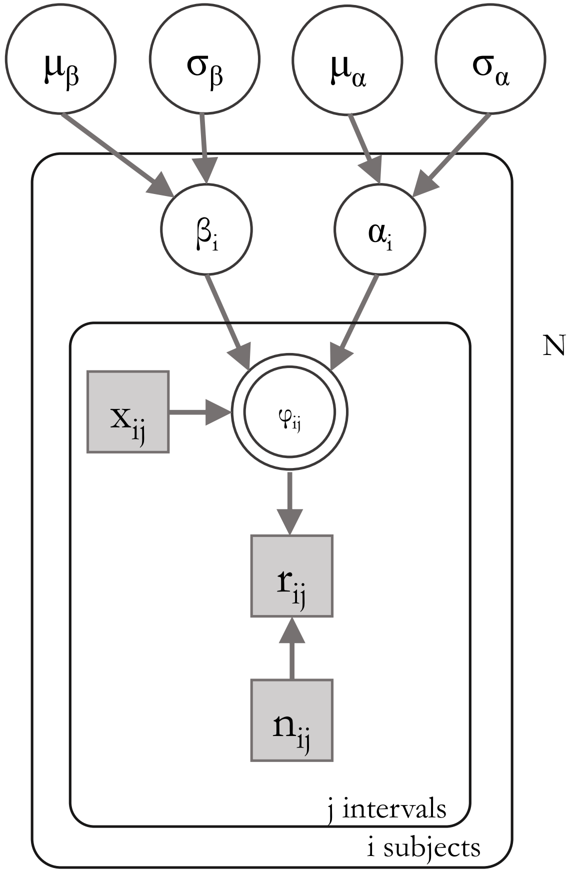
N trials n_i = 5 20 15 ....

Intensities x_i = -20 -10 0 ....

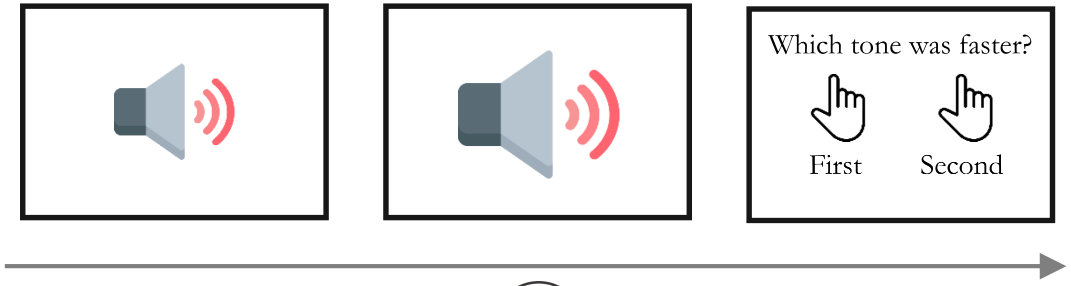


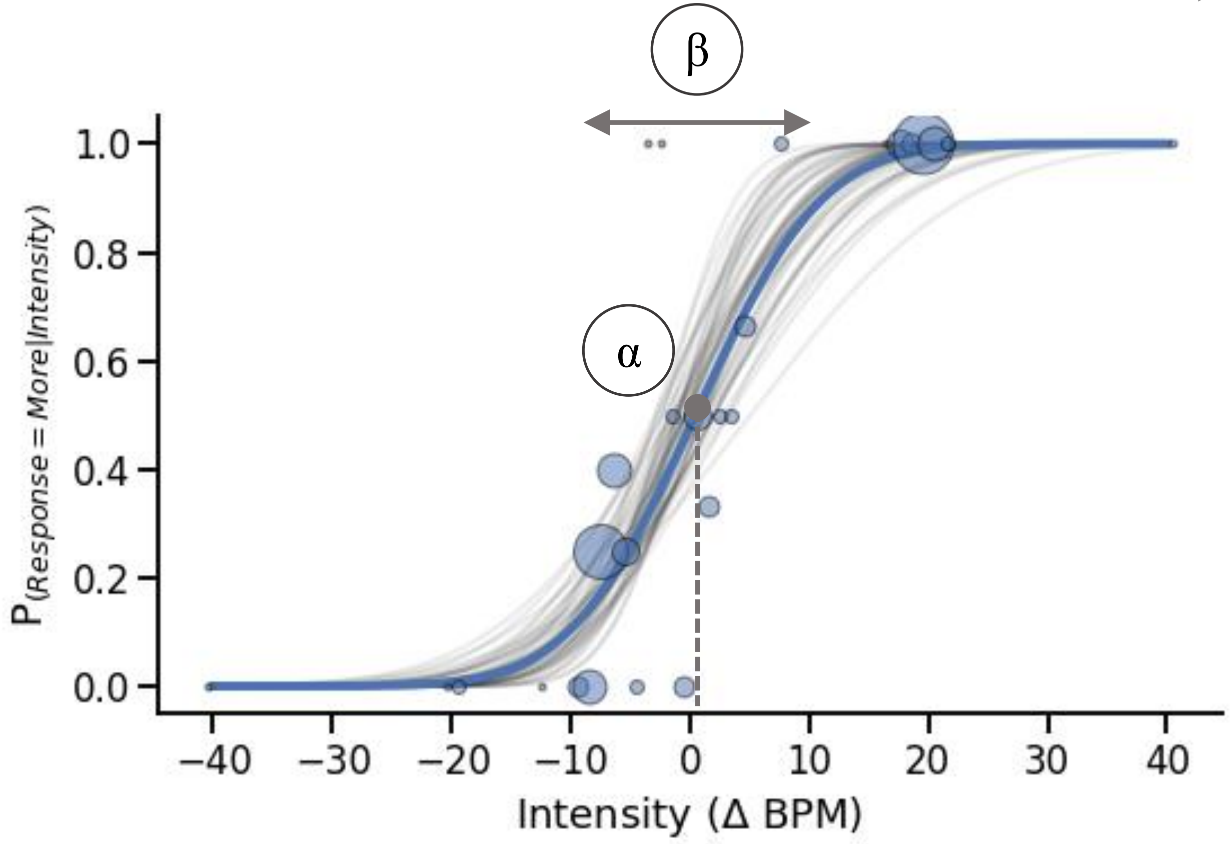
Hierarchial psychophysics

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



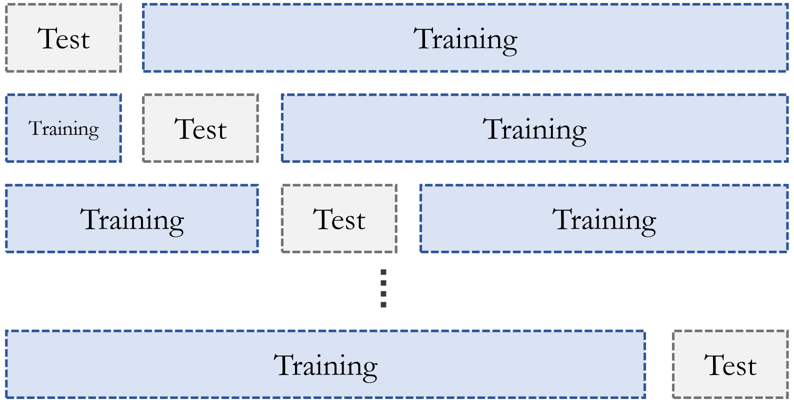
N responses	r_i	=	0	3	10	....
N trials	n_i	=	5	20	15	....
Intensities	x_i	=	-20	-10	0	....
Subject ID	S_i	=	0	0	0	....





Model comparison

Cross-validation



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.

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**

- Widely applicable information Criterion

$$WAIC = -2\mathrm{lppd} + 2\mathrm{p}_{\mathrm{WAIC}}$$

How well the model fits the data **Penalizes complex models**

The generalized linear model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

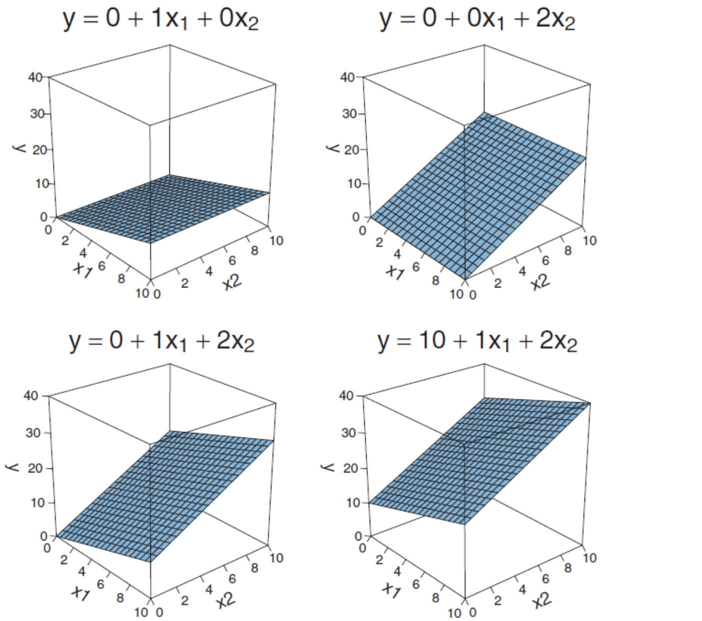


Figure 15.2 Examples of linear functions of two variables, x_1 and x_2. Upper left: Only x_1 has an influence on y. Upper right: Only x_2 has an influence on y. Lower left: x_1 and x_2 have an additive influence on y. Lower right: Nonzero intercept is added.

Interaction term

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \boxed{\beta_3 x_1 x_2}$$

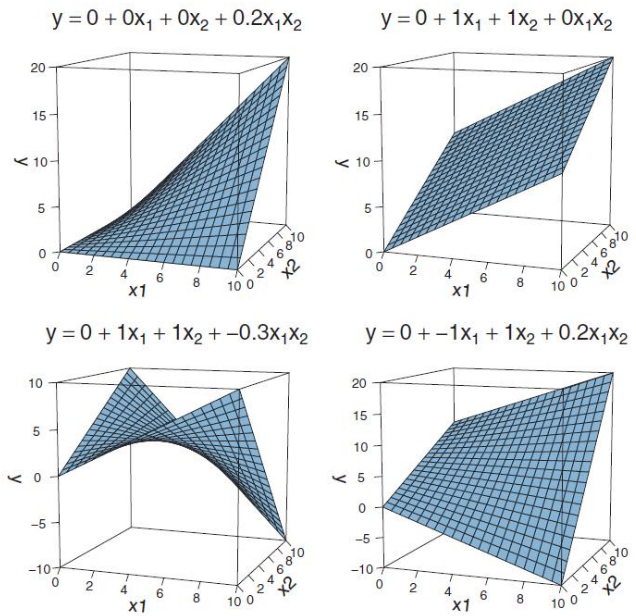


Figure 15.3 Multiplicative interaction of two variables, x_1 and x_2. Upper right panel shows *zero* interaction, for comparison. Figure 18.8, p. 526, provides additional perspective and insight.

$$y = logisitic(x) = \frac{1}{(1+e^{-x})}$$

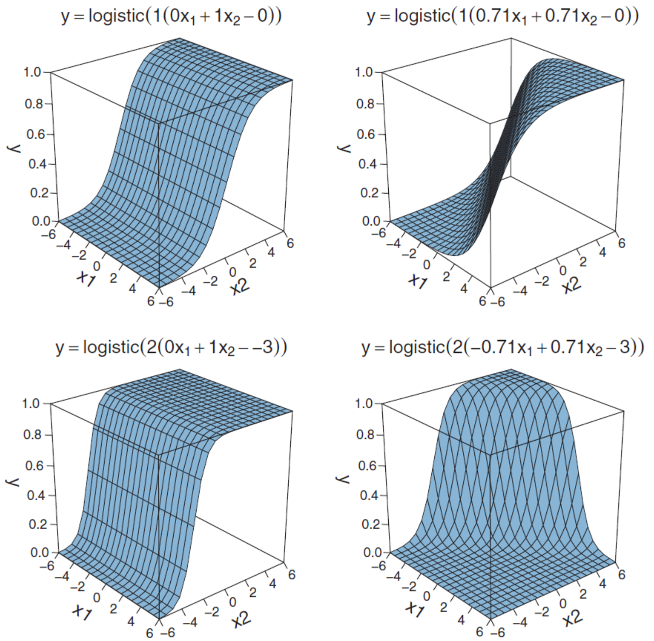


Figure 15.7 Examples of logistic functions of two variables. Top two panels show logistics with the same gain and threshold, but different coefficients on the predictors. The left two panels show logistics with the same coefficients on the predictors, but different gains and thresholds. The lower right panel shows a case with a negative coefficient on the first predictor.

Kruschke, J. (2015). Doing Bayesian data analysis : a tutorial with R, JAGS, and Stan. Boston: Academic Press. Chapter 15.

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.