Fitting RL models to behavior

Tutorial for BAMB 2024

18th July 2024
Nisheet Patel

References included in a pdf, but special credit to Luigi Acerbi:
a long-time friend, mentor, and collaborator

On models

A model is a lie that helps you see the truth.

Howard Skipper

The purpose of models is not to fit the data but to sharpen the questions.

Samuel Karlin

Outline

- Introduction
  - Descriptive vs. Process models
  - Psychometric function and RL models and their likelihood functions
  - Example task: 2AFC
- Model fitting, selection, and recovery
  - As a statistical estimation problem
  - Optimizing the likelihood function
  - Tips for model selection for sequential data
  - Basics of model recovery
- A cautionary tale about model fitting
  - Good experimental design and why steady-state behavior matters

Outline

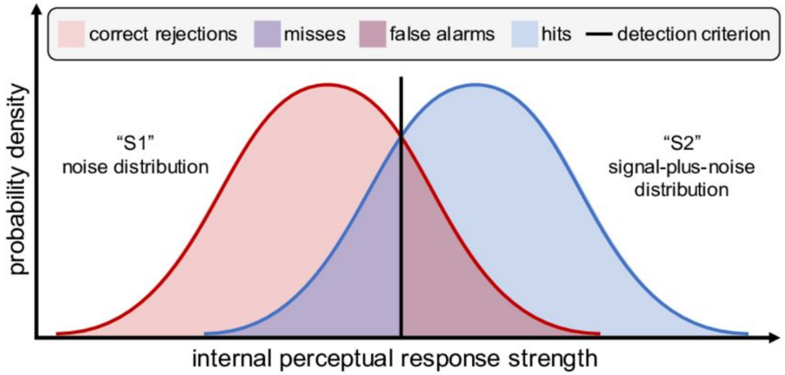
- Introduction
  - Descriptive vs. Process models
  - Psychometric function and RL models and their likelihood functions
  - Example task: 2AFC
- Model fitting, selection, and recovery
  - As a statistical estimation problem
  - Optimizing the likelihood function
  - Tips for model selection for sequential data
  - Basics of model recovery
- A cautionary tale about model fitting
  - Good experimental design and why steady-state behavior matters

Two kinds of models

- Descriptive model
  - Purpose:
    - Describes observed relationships in the data
  - Examples:
    - The psychometric function
    - neural networks
- Process model
  - Purpose:
    - Seeks to explain the mechanisms underlying psychological phenomena
  - Examples:
    - Signal Detection Theory
    - Drift-diffusion models
    - Bayesian (ideal observers)
    - Reinforcement Learning models

Two kinds of models

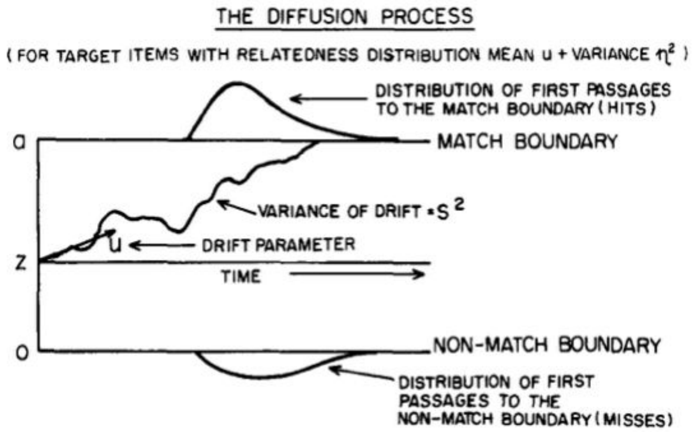
- Descriptive model
  - Purpose:
    - Describes observed relationshi
  - Examples:
    - The psychometric function
    - neural networks
- Process model
  - Purpose:
    - Seeks to explain the mechanisms underlying psychological phenomena
  - Examples:
    - Signal Detection Theory
    - Drift-diffusion models
    - Bayesian (ideal observers)
    - Reinforcement Learning models



Fechner, Heeger, et al.; Fig: Morales 2021

Two kinds of models

- Descriptive model
  - Purpose:
    - Describes observed relationships ir
  - Examples:
    - The psychometric function
    - neural networks
- Process model
  - Purpose:
    - Seeks to explain the mechanisms underlying psychological phenomena
  - Examples:
    - Signal Detection Theory
    - Drift-diffusion models
    - Bayesian (ideal observers)
    - Reinforcement Learning models



Ratcliff 1978: A theory of memory retrieval

Two kinds of models

- Descriptive model
  - Purpose:
    - Describes observ
  - Examples:
    - The psychometr
    - neural networks
- Process model
  - Purpose:
    - Seeks to explain the mechanisms underlying psychological phenomena
  - Examples:
    - Signal Detection Theory
    - Drift-diffusion models
    - Bayesian (ideal observers)
    - Reinforcement Learning models

Bayesian models:

- Special kind of process models
- Reason about sources of uncertainty
  - Noise
  - Ambiguity

e.g. Drugowitsch et al 2016; RL as inference: Levine 2018; SQL: Haarnoja et al. 2017-18

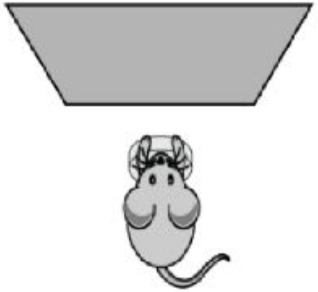
International Brain Laboratory

Experimental & theoretical neuroscientists collaborating to understand brainwide circuits for complex behavior

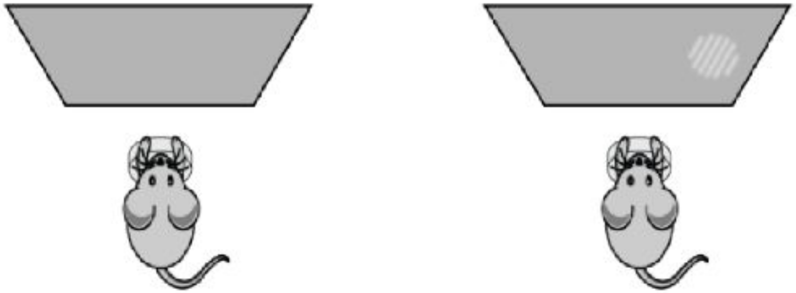
The International Brain Laboratory will release all data sets within 12 months of collection, or upon acceptance for publication of an associated manuscript, whichever comes first.

We will use their task as a guiding example

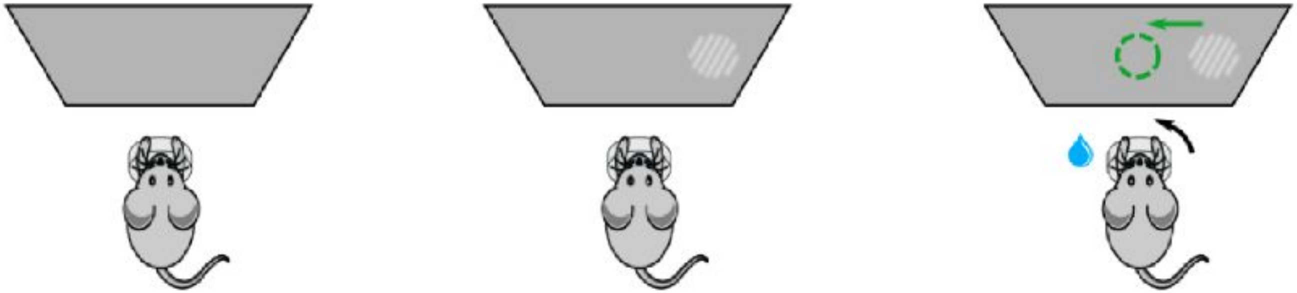
The IBL task



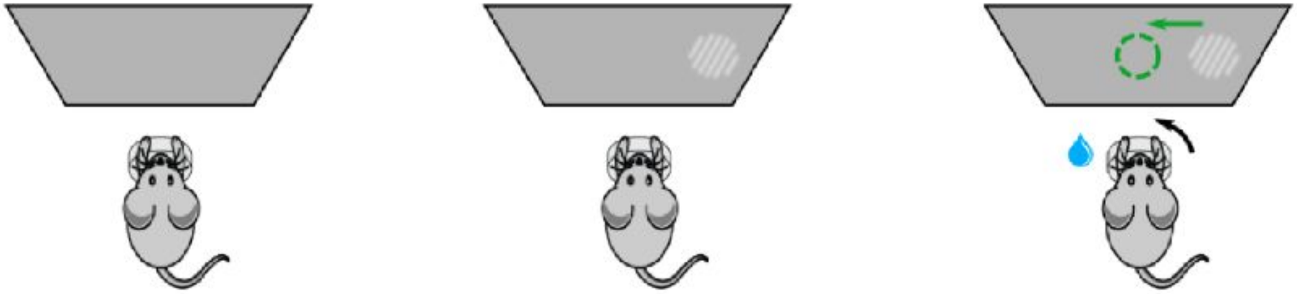
The IBL task

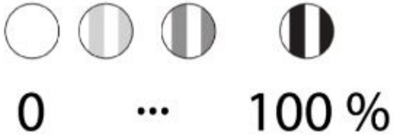


The IBL task

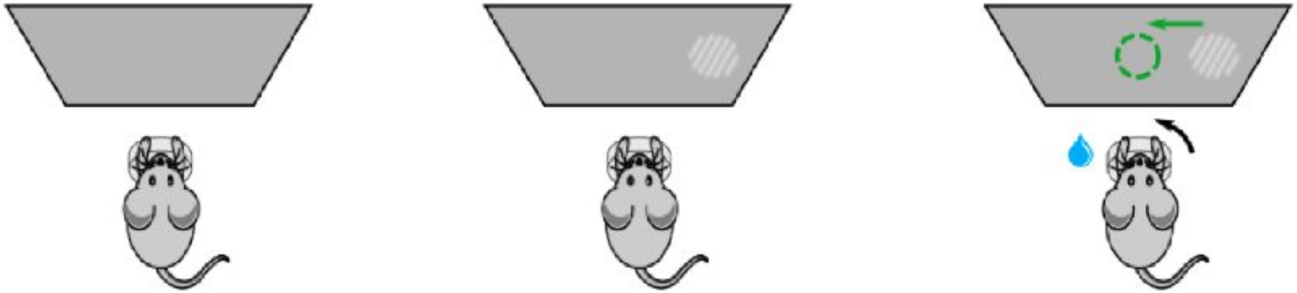


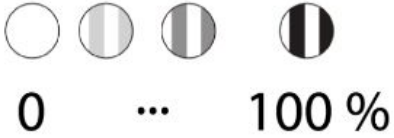
The IBL task

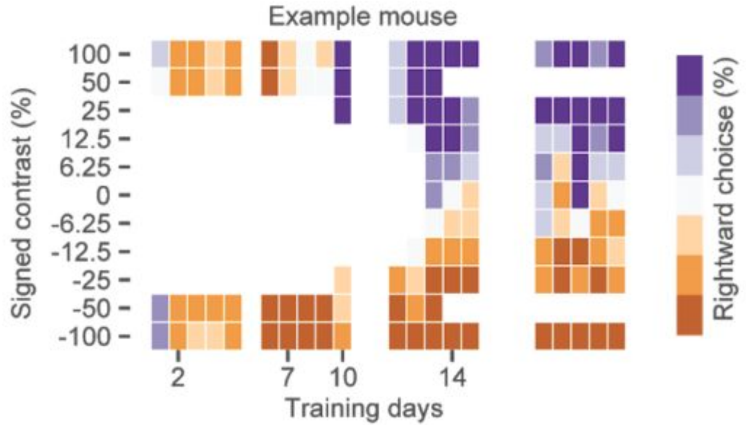




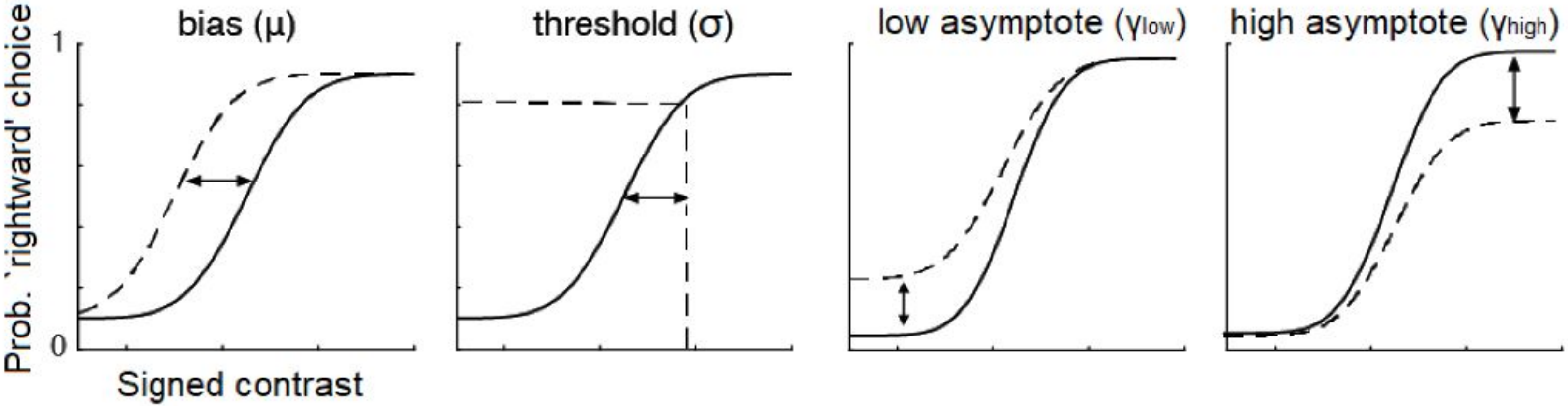
The IBL task




0
...
100 %


Example mouse
Signed contrast (%)
100
50
25
12.5
6.25
0
-6.25
-12.5
-25
-50
-100
2
7
10
14
Training days
Rightward choice (%)

The psychometric function as a descriptive model



- Data: (signed contrast, choice) for each trial
- Parameters θ: $(\mu, \sigma, \gamma_{\text{low}}, \gamma_{\text{high}})$

$$p(\text{rightward choice} | s, \theta) = \gamma_{\text{low}} + (1 - \gamma_{\text{high}} - \gamma_{\text{low}}) \cdot F(s; \mu, \sigma)$$

The (log) likelihood

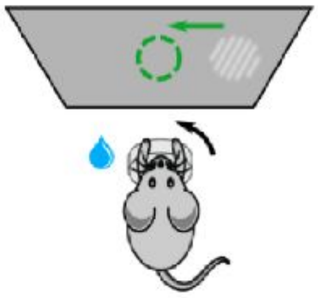
- p(data|θ) is a *probability density* as you vary data for fixed θ
- p(data|θ) is the *likelihood*, a function of θ for fixed data

- For numerical reasons, we work with log p(data|θ)

- For descriptive models that ignore the sequential nature of the data:

$$\log p(\text{data}|\theta) = \sum_{i=1}^{N} \log p_i\left(r^{(i)} \middle| s^{(i)}, \theta\right)$$

A Reinforcement Learning model



- Learn values with TD-learning, here, Q-learning

$$Q(s_t, a_t) = (1 - \alpha)Q(s_t, a_t) + \alpha \cdot \delta$$

$$\delta = r(s_t, a_t) + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)$$

- Use a softmax choice policy

$$\pi(a|s_t) \propto \exp(\beta \cdot Q(s_t, a))$$

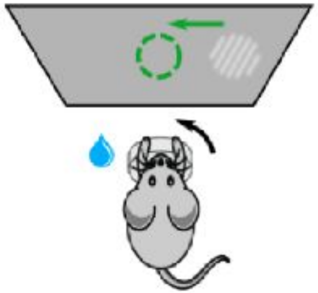
$$\pi(a|s_t) = \frac{\exp(\beta \cdot Q(s_t, a))}{1 + \exp(\beta \cdot Q(s_t, a))}$$

Note: this makes several implicit assumptions.

In fact, it optimizes a slightly different objective specified by maximum entropy RL (see Levine 2018: RL as inference)

A Reinforcement Learning model

- Data: (signed contrast, choice) for each trial
- Parameters θ: (α, β)

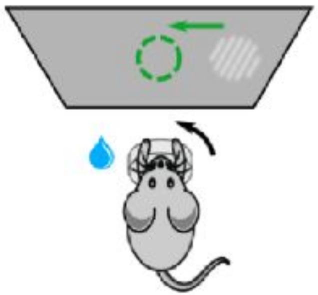


$$p(a|s;\theta) = \pi(a_t|o_t, M, o_{1:t-1}, a_{1:t-1}; \theta_t)$$

- Given the data, this is a deterministic function of the parameters!

- For RL algorithms with stochastic update rules, e.g. policy gradient methods using SGD, this likelihood would be noisy
(but sometimes there is no way around it)

A Reinforcement Learning model



- Data: (signed contrast, choice) for each trial
- Parameters θ: (α, β)

$$p(a_t | s_t; \theta) = \pi(a_t | o_t, M, o_{1:t-1}, a_{1:t-1}; \theta)$$

- With the likelihood for one trial, we can compute it sequentially for the data:

$$\log p(\text{data} | \theta) = \sum_{t=1}^{T} \log \pi(a_t | o_t, M, o_{1:t-1}, a_{1:t-1}; \theta)$$

Outline

- Introduction
  - Descriptive vs. Process models
  - Psychometric function and RL models and their likelihood functions
  - Example task: 2AFC
- Model fitting, selection, and recovery
  - As a statistical estimation problem
  - Optimizing the likelihood function
  - Tips for model selection for sequential data
  - Basics of model recovery
- A cautionary tale about model fitting
  - Good experimental design and why steady-state behavior matters

Model fitting

as a statistical estimation problem

1. Maximum likelihood estimation (MLE)
    - Find the maximum of p(data|θ)

$$\hat{\theta}_{ML} = \arg\max_{\theta} p(\text{data}|\theta) = \arg\max_{\theta} \log p(\text{data}|\theta)$$

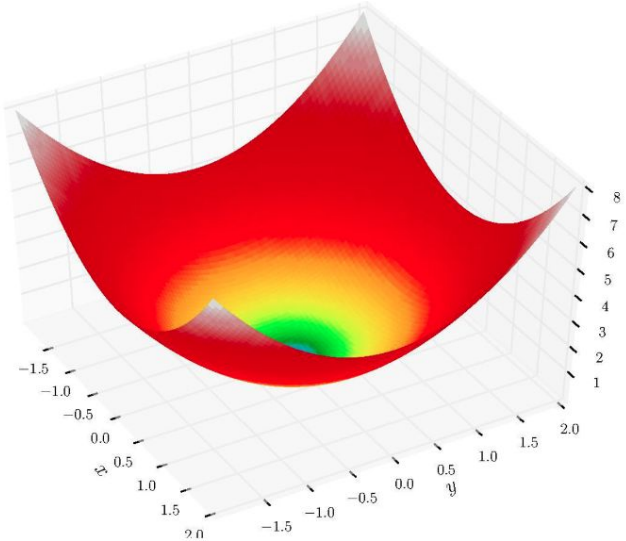
2. Bayesian posterior

$$p(\theta|\text{data}) \propto p(\text{data}|\theta)p(\theta)$$

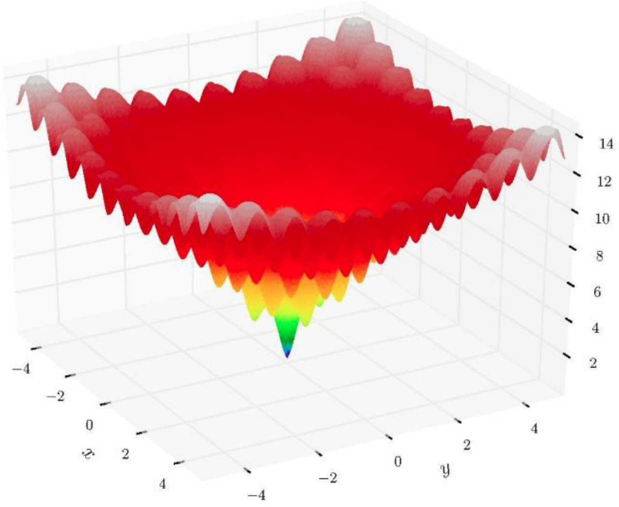
- Full posterior is informative about parameter uncertainty
- But computationally expensive

Now we can fit the models!

Expectation



Reality

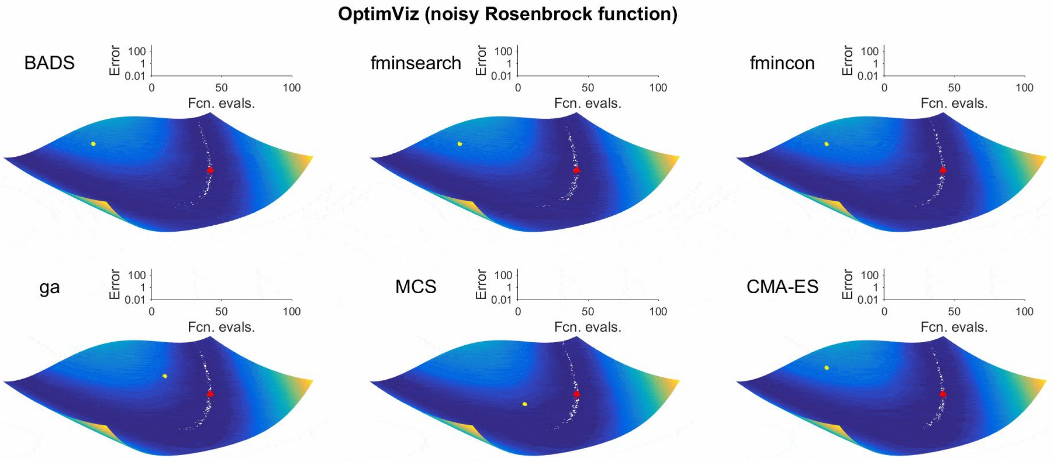


Model fitting is hard!

Toolboxes for fitting RL models (with noisy likelihoods)

Bayesian Adaptive Direct Search

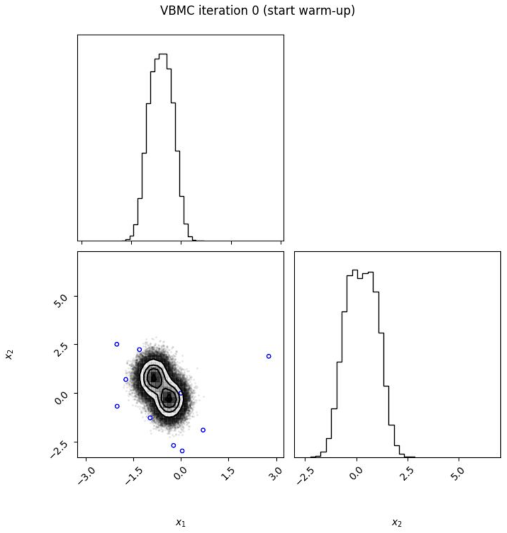
(≈ Direct Search + Bayesian Optimization)



Acerbi & Ma 2017

Variational Bayesian Monte Carlo

(≈ Variational inference + Monte Carlo)

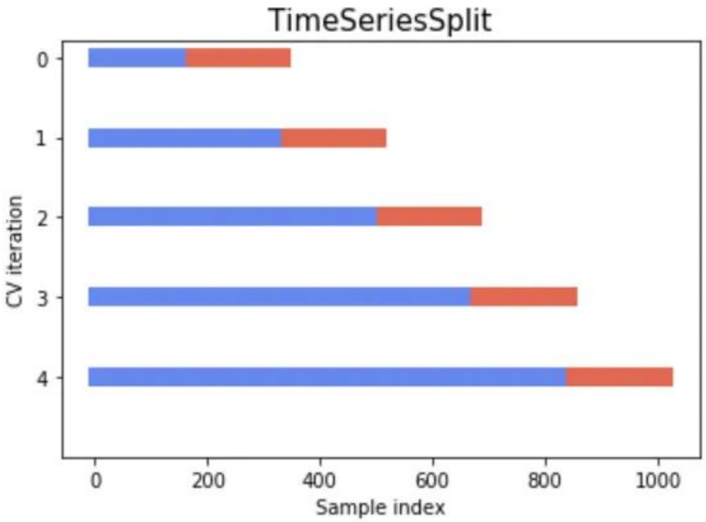


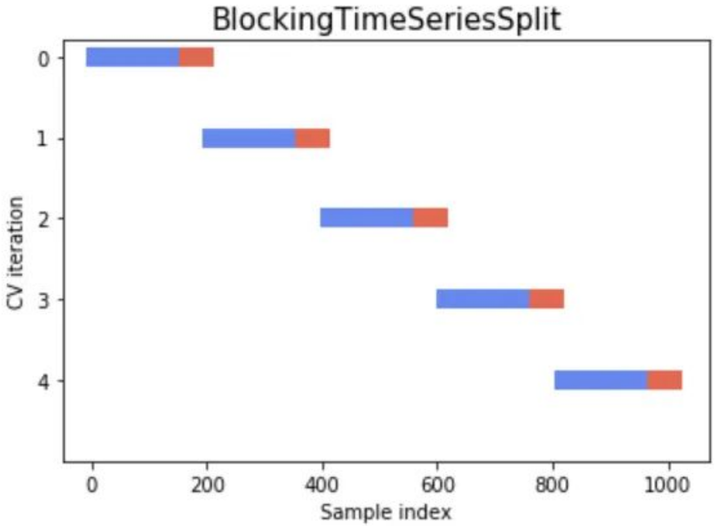
Acerbi 2018

www.github.com/acerbilab

Tips for model selection for RL models

- Be careful with cross validation - data is non iid





from scipy.stats import TimeSeriesSplit, BlockingTimeSeriesSplit

Some additional tips and resources

- Wilson & Collins 2019:
    - Ten simple rules for computational modeling for behavioral data
- Danwitz et al. 2021:
    - Parameter and model recovery of reinforcement learning models
- Barbosa et al. 2023:
    - A practical guide for studying human behavior in the lab

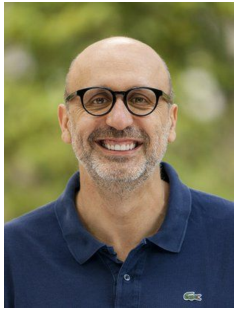
Outline

- Introduction
  - Descriptive vs. Process models
  - Psychometric function and RL models and their likelihood functions
  - Example task: 2AFC
- Model fitting, selection, and recovery
  - As a statistical estimation problem
  - Optimizing the likelihood function
  - Tips for model selection for sequential data
  - Basics of model recovery
- A cautionary tale about model fitting
  - Good experimental design and why steady-state behavior matters

On model fitting

Claims about behavior that are solely based on results from model fitting are hard to believe. I like to design experiments such that the effects we're after are obnoxiously staring in your face.
I call it the Ocular Obviousness Test.

Antonio Rangel



Rec.: Optimal Experimental Design

Questions?