Exploring sequential sampling models of behaviour

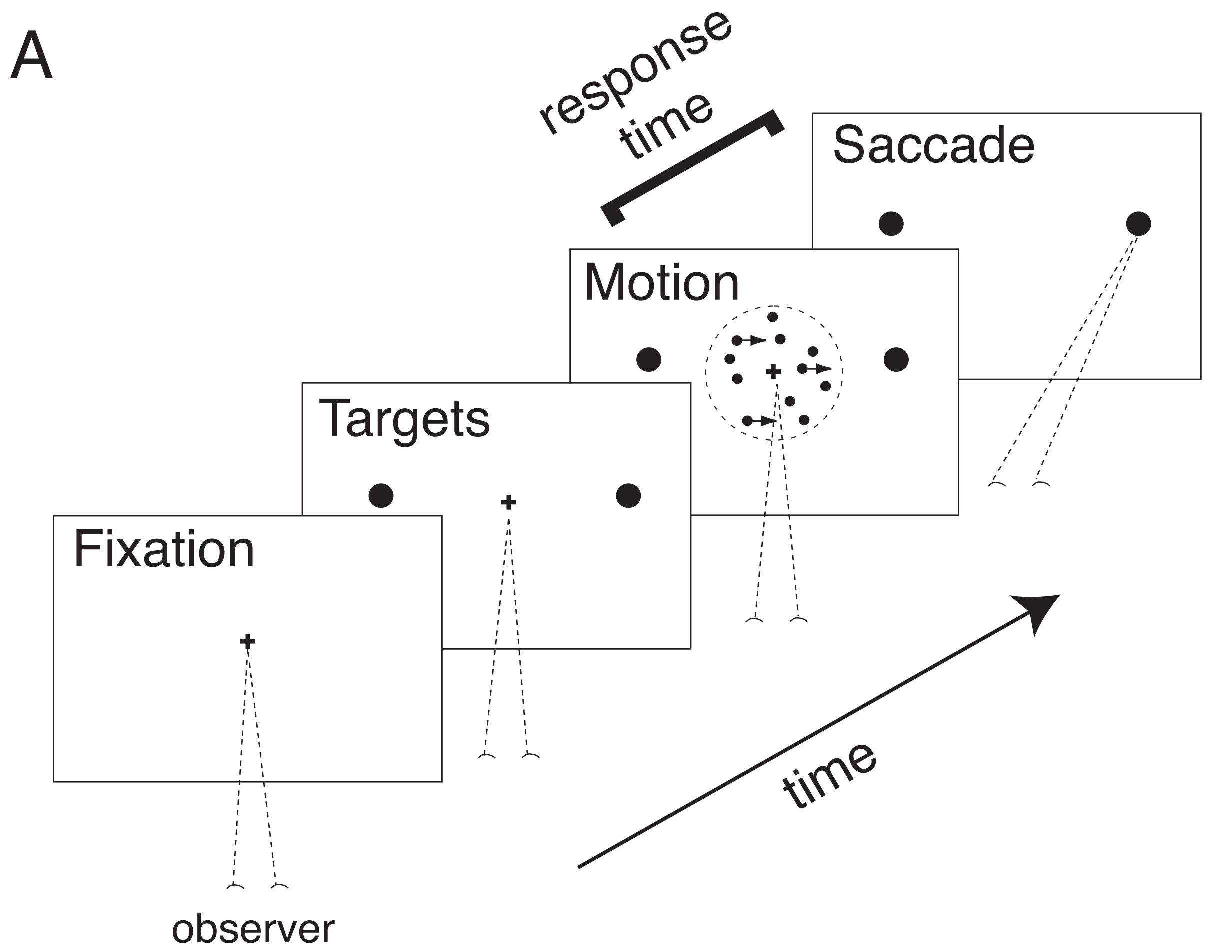
Guy Hawkins
guy.e.hawkins@gmail.com
University of Amsterdam

What should you get out of today?

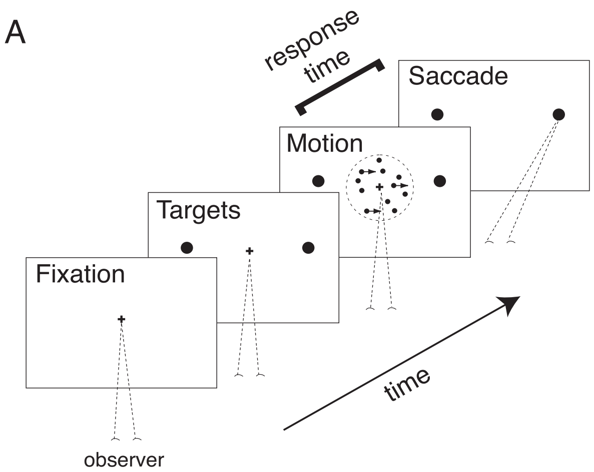
1. Describe the connection between psychological processes, model parameters, and behaviour in the context of decision making

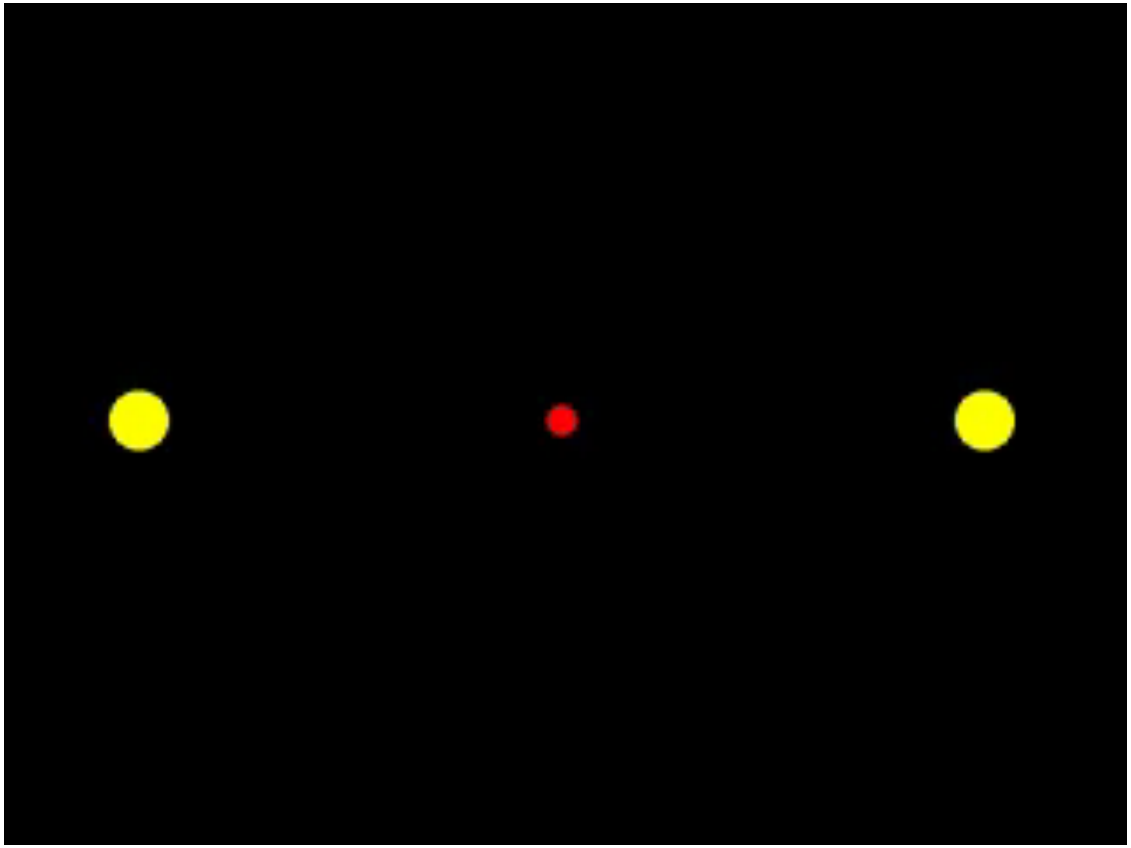
2. Conceptual understanding of the processes involved in setting up a model, fitting it to data, and evaluating the quality of model fit to data (visual, model selection index)

2AFC task


A
response
time
Saccade
Motion
Targets
Fixation
time
observer

2AFC task


A
response time
Saccade
Motion
Targets
Fixation
observer
time



Why Model Decisions?

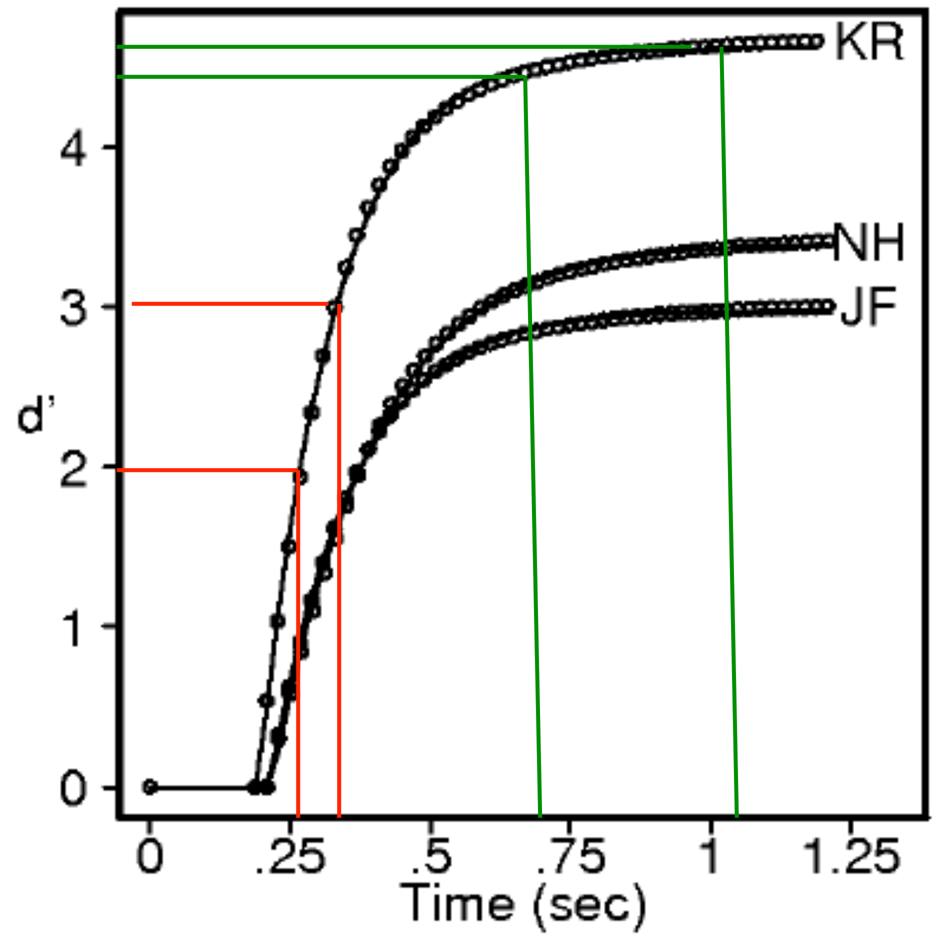
Evidence Accumulation Model (EAM)

MANIFEST = Response & *Distribution* of Response Times (RT)

EAM

LATENT = e.g.,
quality of evidence,
caution (evidence *threshold*),
non-decision time,
…

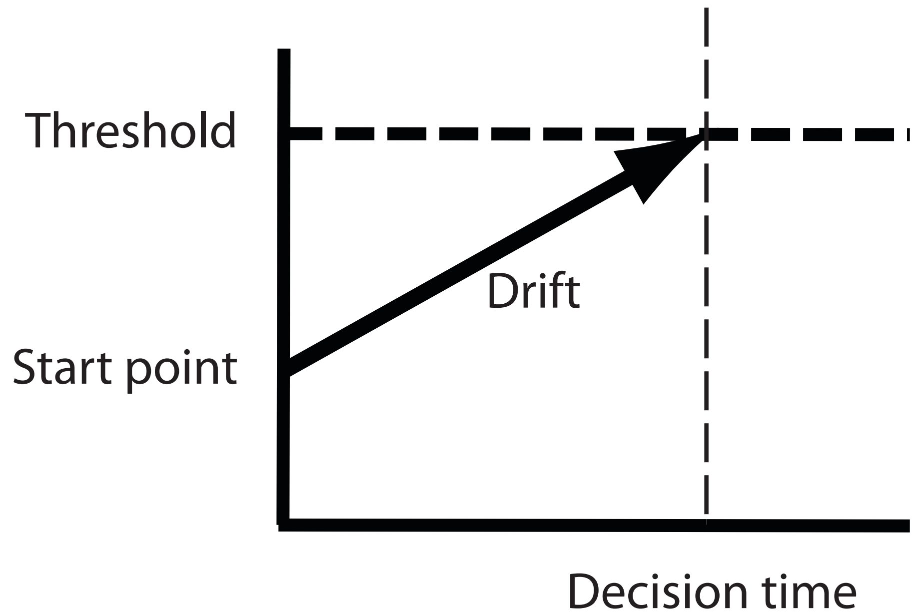
Speed-Accuracy Tradeoff

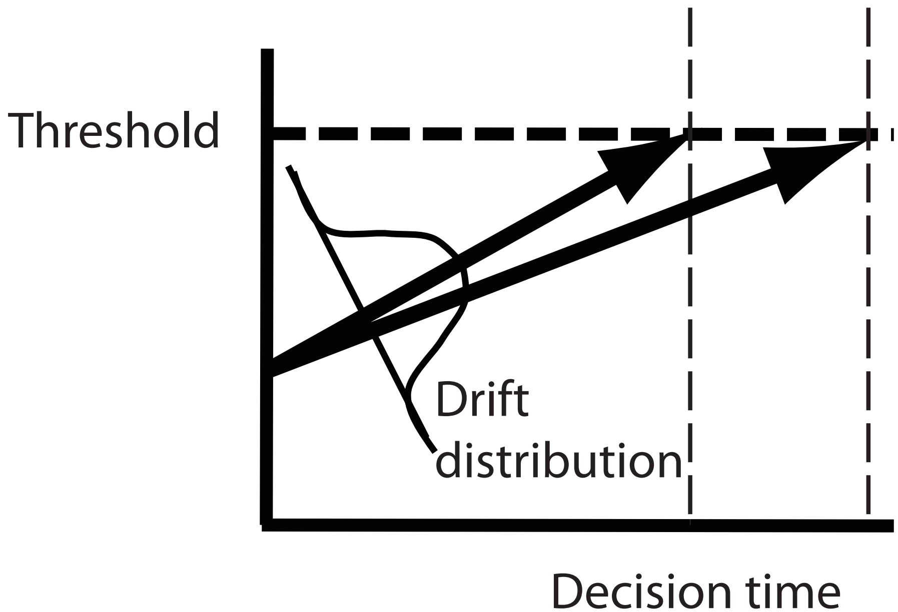


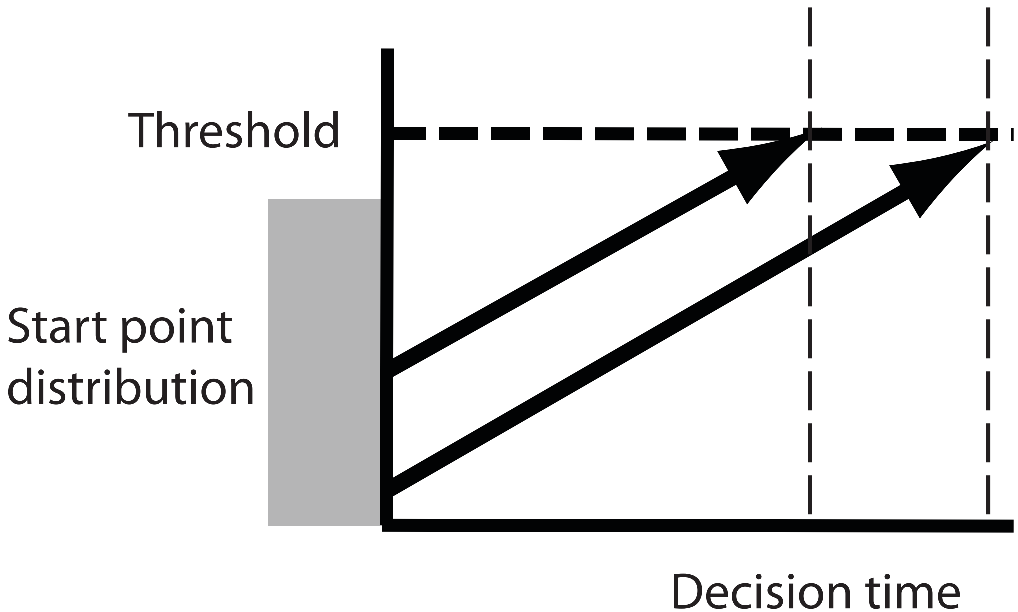
Why quantitative models?

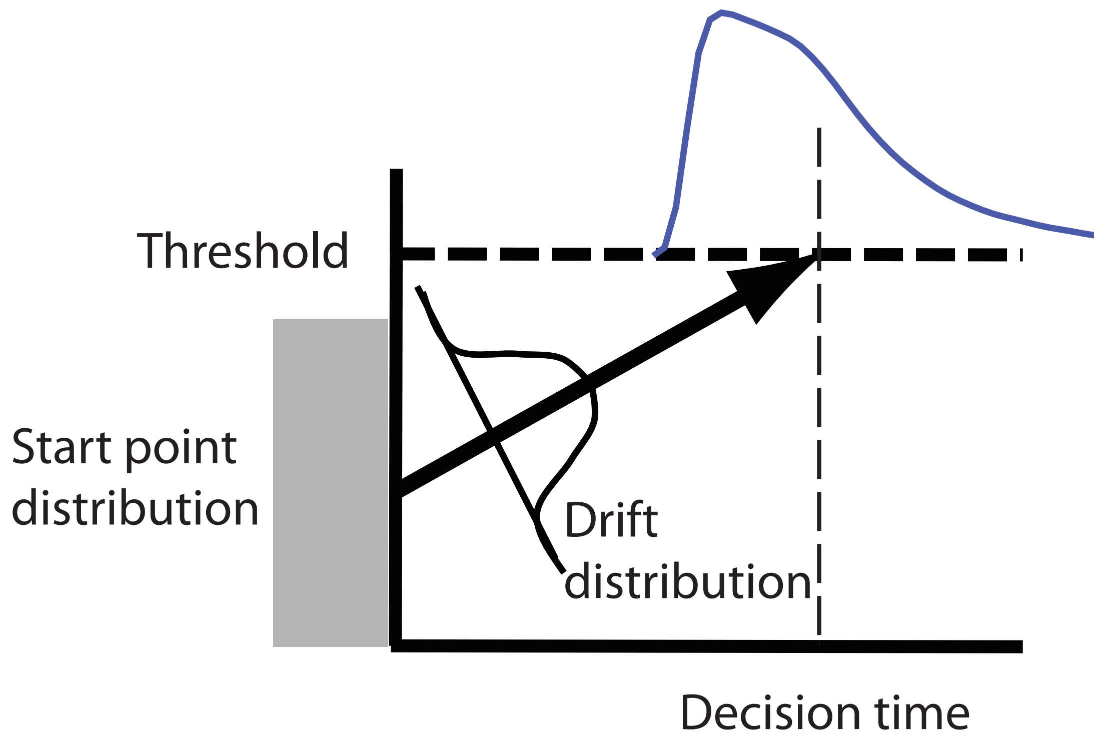
- Cognitive process models are quantitative implementations of theories about the processes involved in cognition
  - Memory, attention, language, decision-making…
- Decompose observed variables into latent components of processing
  - Observed variables: choices, response times
  - Latent components: response caution, efficiency of processing, non-decision, etc.
- Allow quantification of the evidence for competing (quantitative) theories

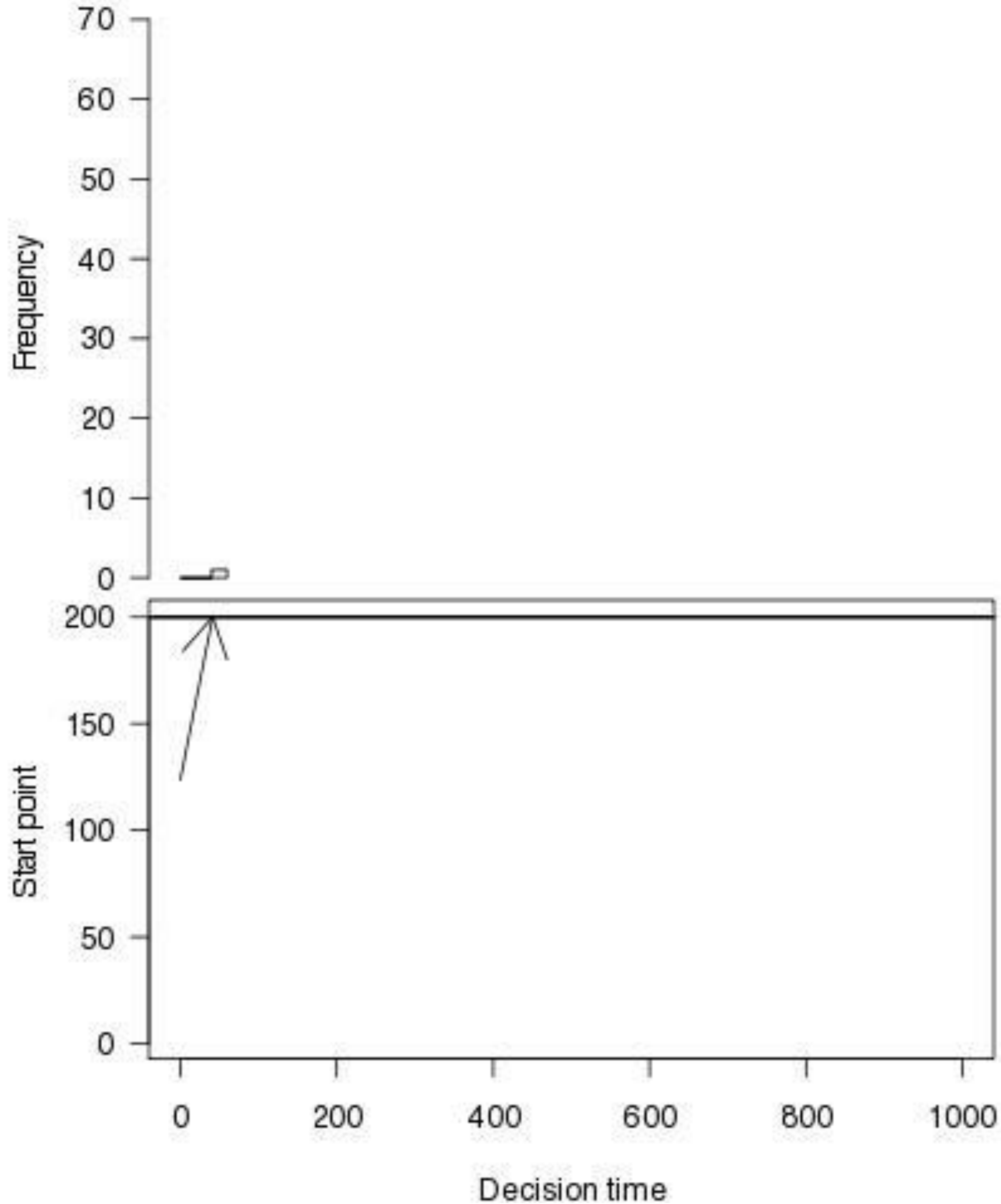
Linear Ballistic Accumulator (LBA)



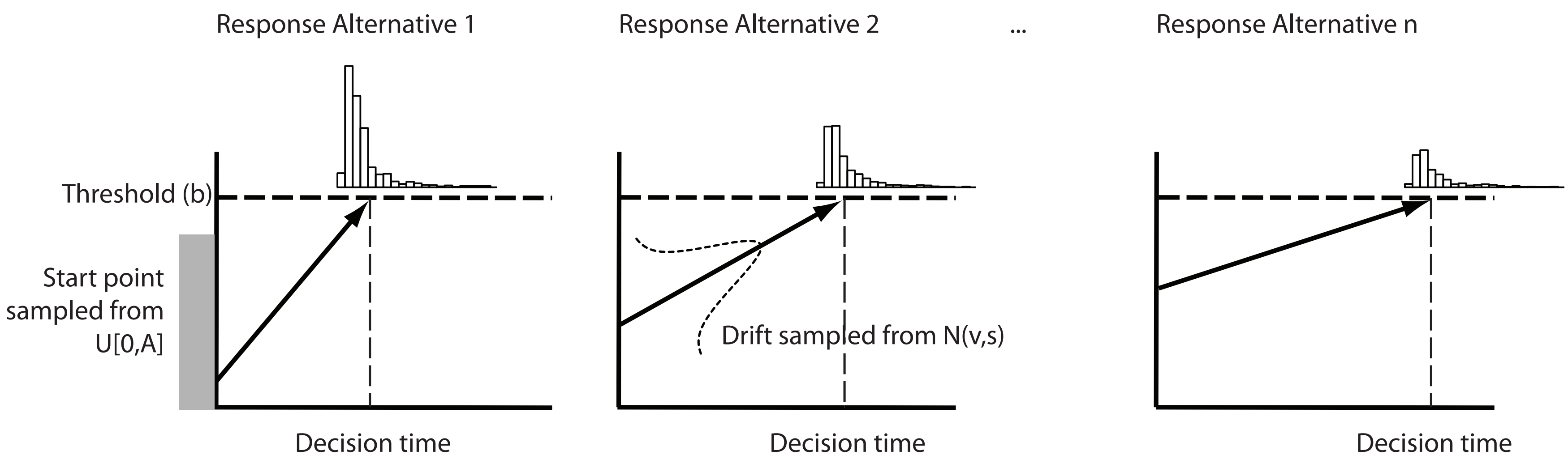

Threshold
Drift
distribution
Decision time


Threshold
Start point
distribution
Decision time


Threshold
Start point distribution
Drift distribution
Decision time

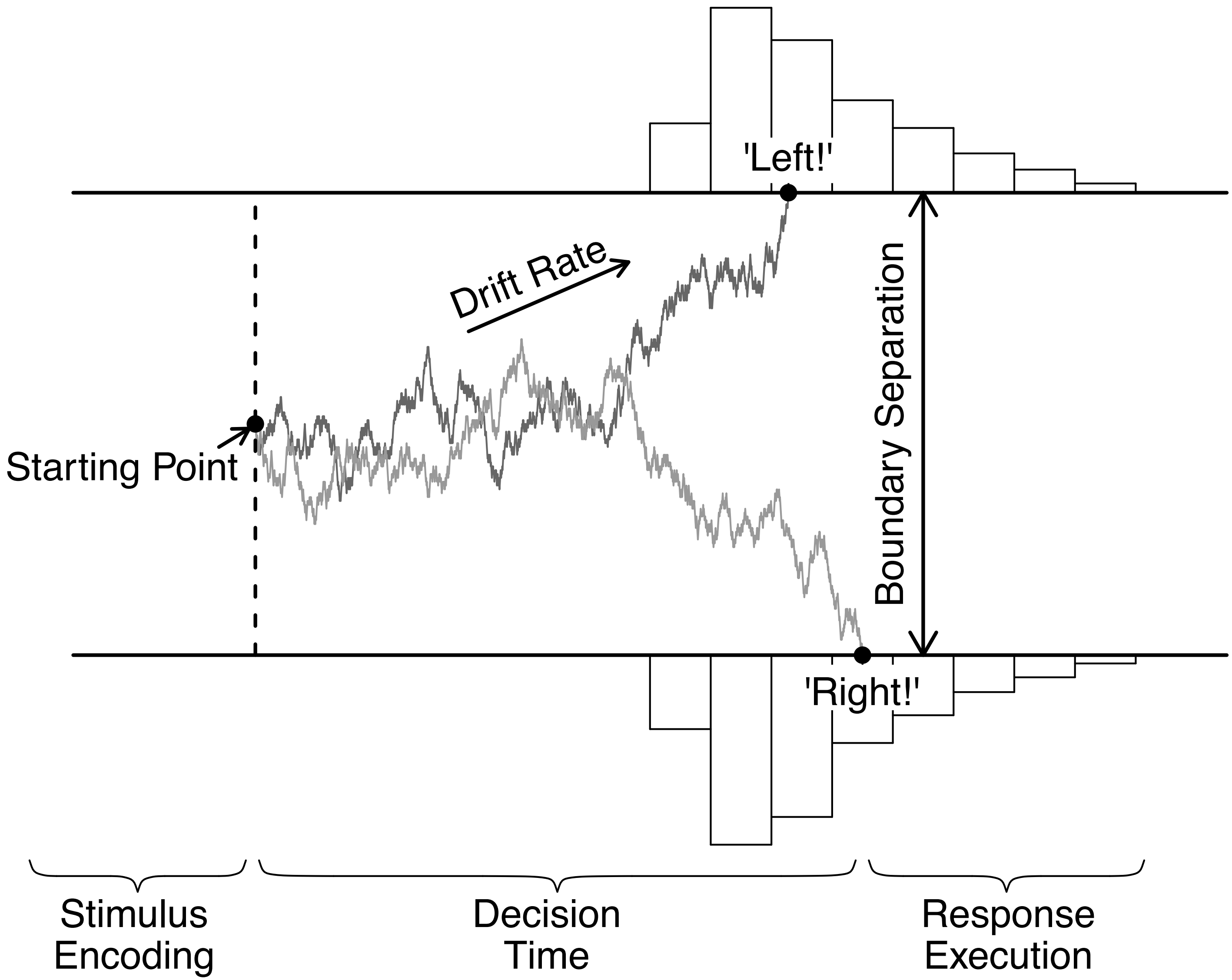

Frequency
70
60
50
40
30
20
10
0
Start point
200
150
100
50
0
0
200
400
600
800
1000
Decision time

Linear Ballistic Accumulator (LBA)



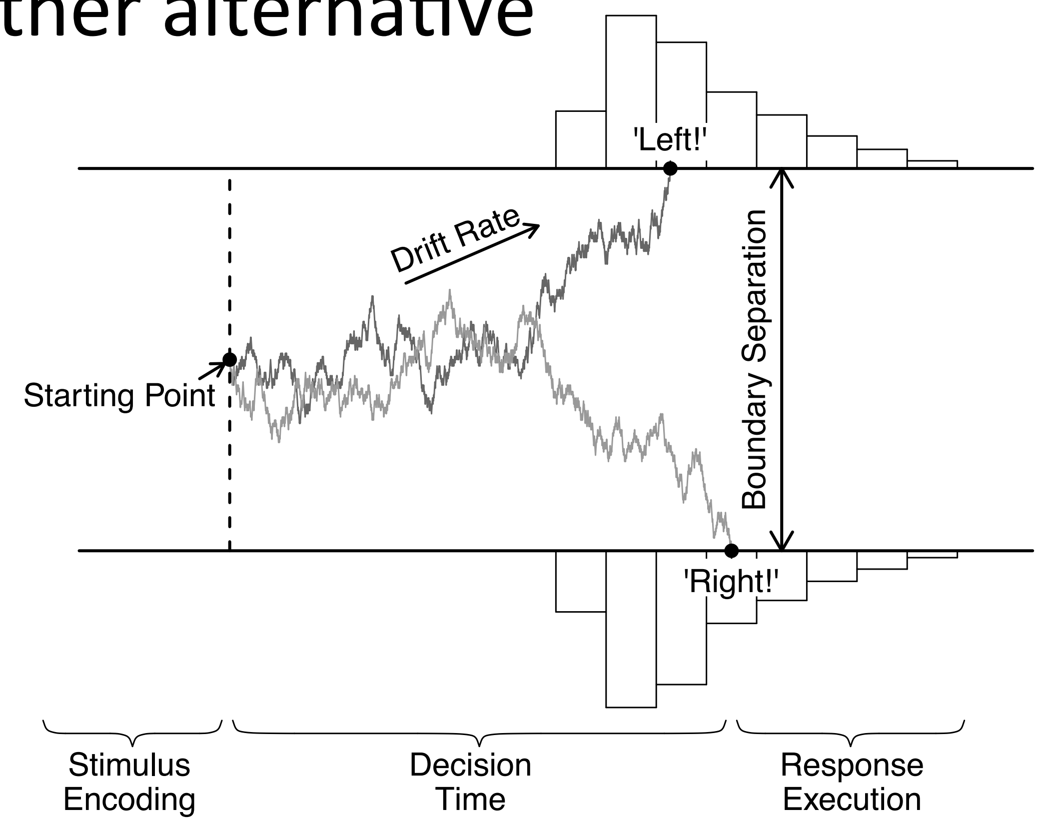
- Each alternative has its own accumulator
- Fastest accumulator is selected

Diffusion Decision Model (DDM)



Diffusion Decision Model (DDM)

- Stochastic
- Evidence in favour of one alternative = evidence against other alternative
- 2-alternative!



The general idea

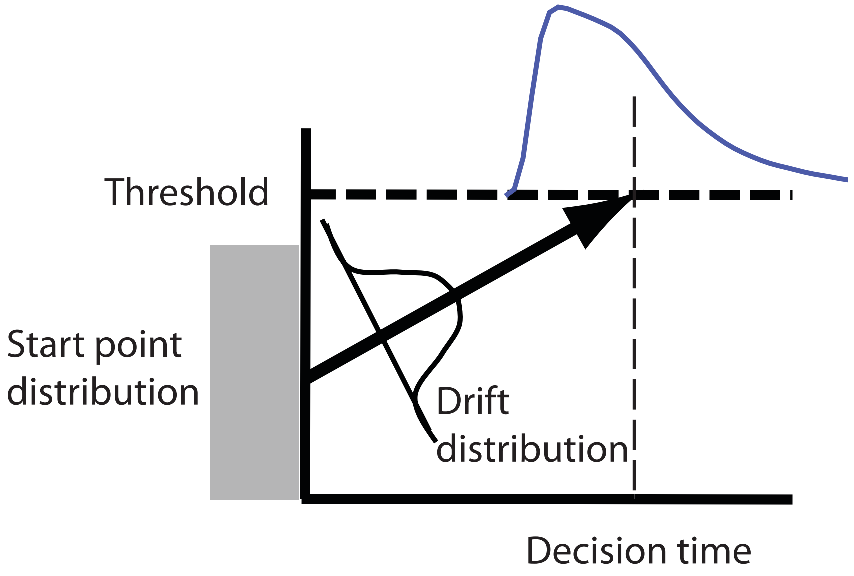
- Sequential sampling models have parameters
  - Parameters have a psychological interpretation
    - Drift rate (processing efficiency), response threshold (caution), start point (bias), etc.

- Different model parameters make different predictions for patterns of response accuracy and the shape of correct and error response time distributions

- Use the observed pattern of response accuracy and response time distributions to infer the most likely combination of parameters to have generated the data
  - Provide a psychological interpretation of the data

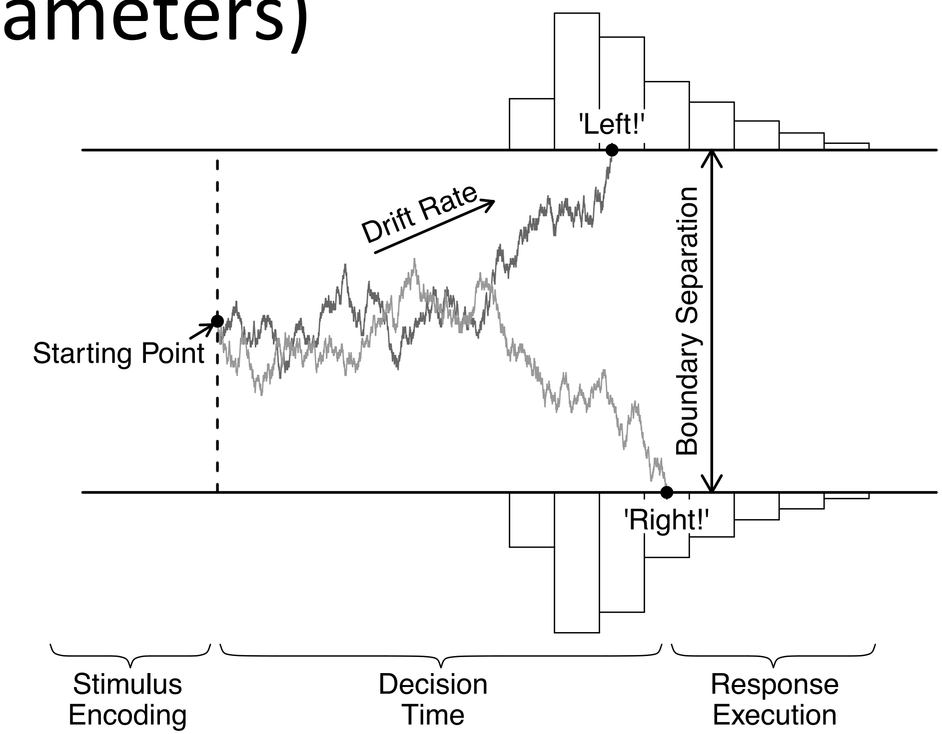
Simulating data from the LBA

- LBA parameters (5 parameters)
  - Start point: U[0,**A**]
  - Threshold: **b**
  - Drift: N(**v**,s), typically
    - Correct drift rate: $\mathbf{v_c}$
    - Error drift rate: $\mathbf{v_e}$
  - Non-decision time: **t0**



Simulating data from the DDM

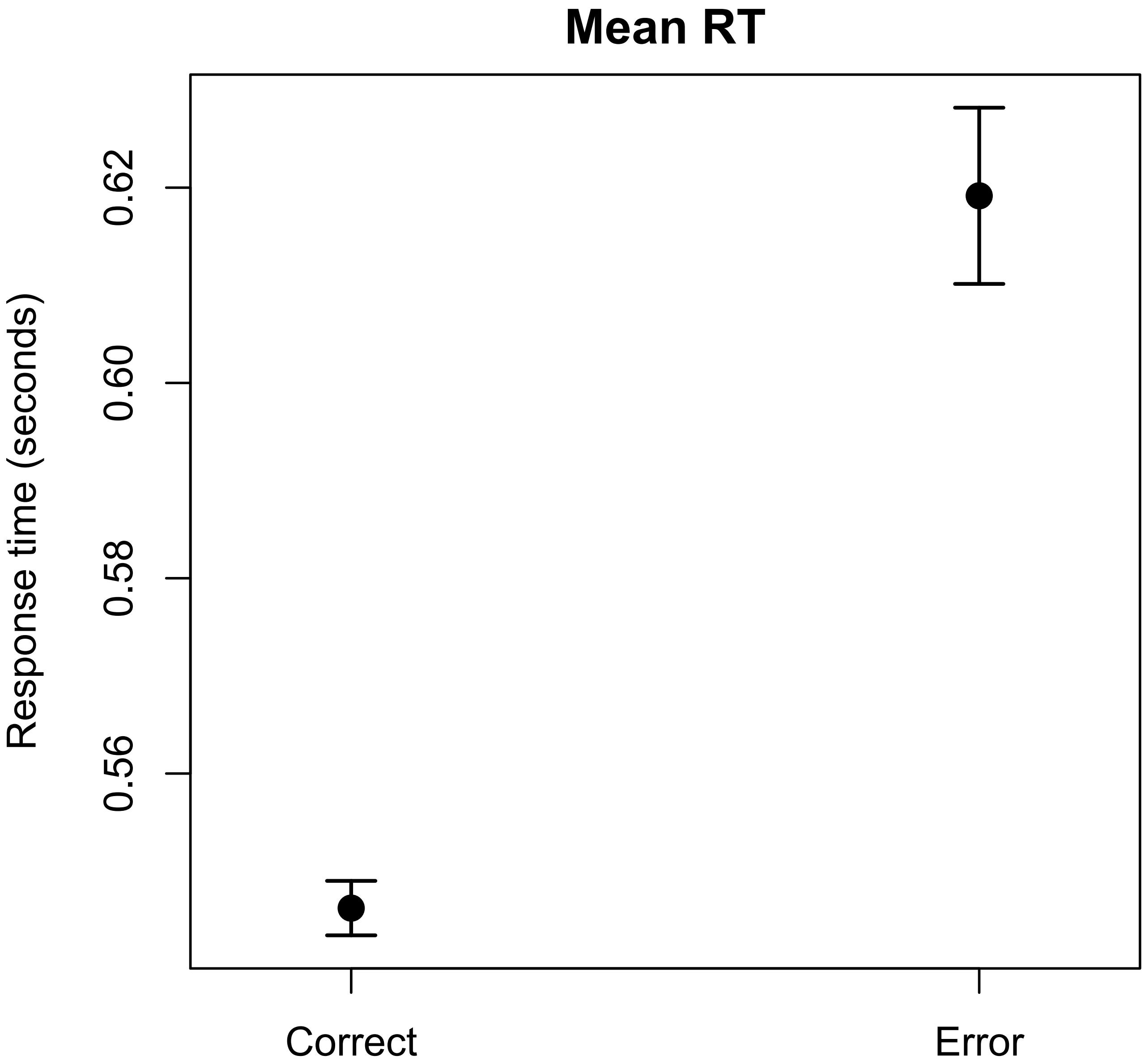
- DDM parameters (7 parameters)
  - Start point:
    - Mean: **z**
    - Range: U[z-**sz**/2, z+**sz**/2]
  - Boundary separation: **a**
  - Drift rate: N(**v**,**η**)
  - Non-decision time
    - Mean: **t0**
    - Range U[t0-**st0**/2, t0+**st0**/2]

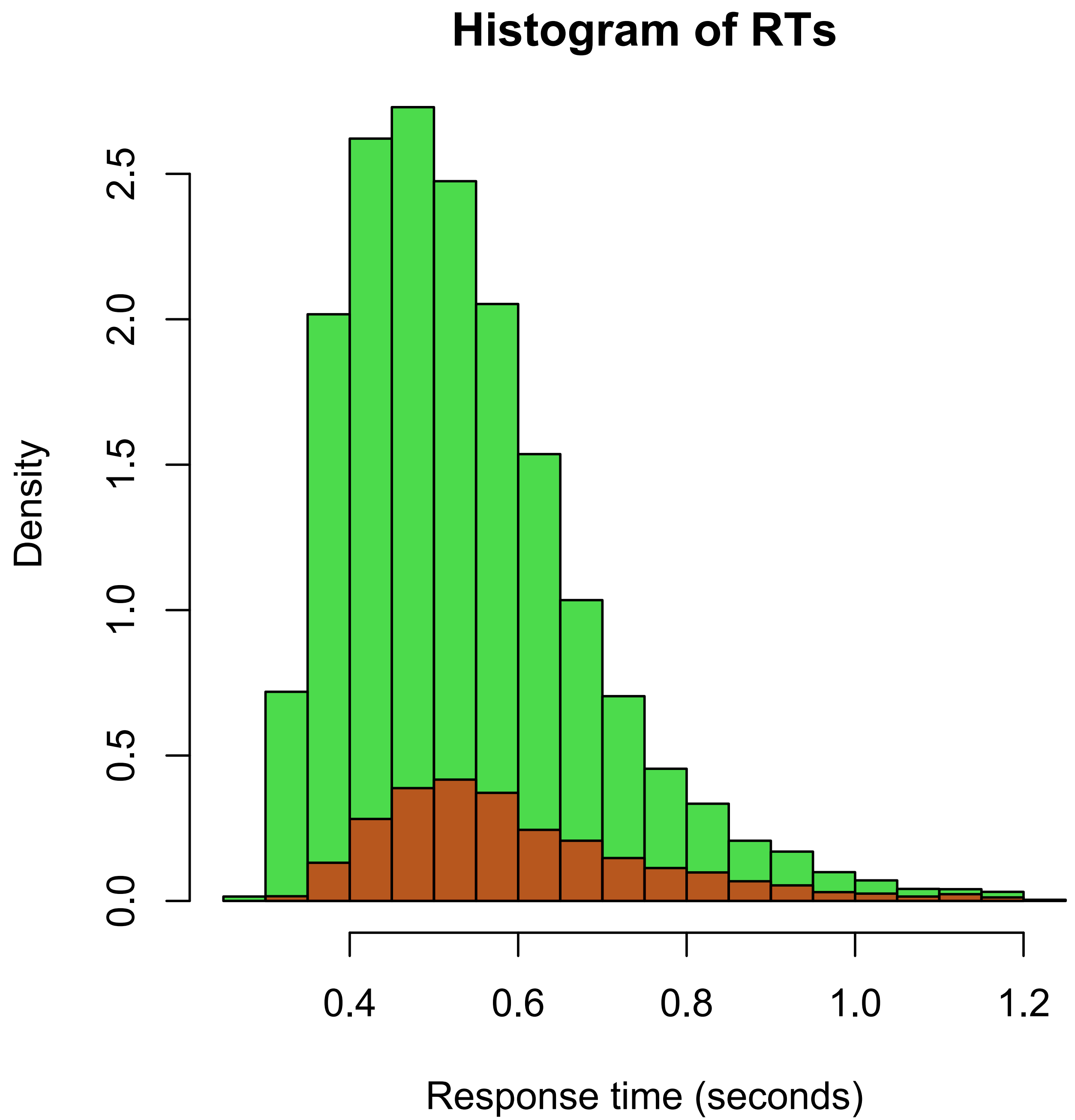


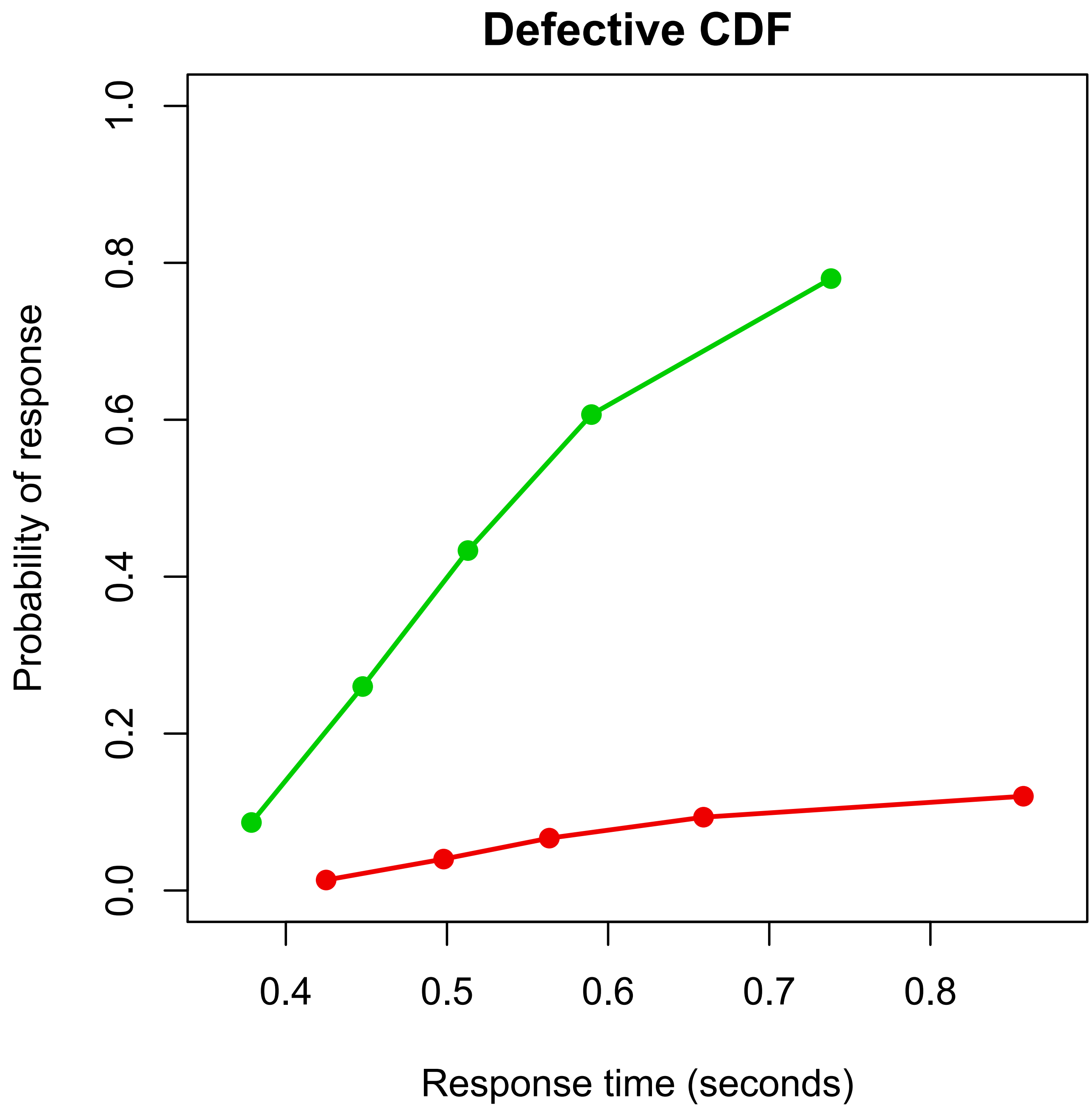
Model parameters

- One by one, we'll step through the effect of changing model parameters on predicted behaviour
  - We'll only cover the effect of "core" model parameters today

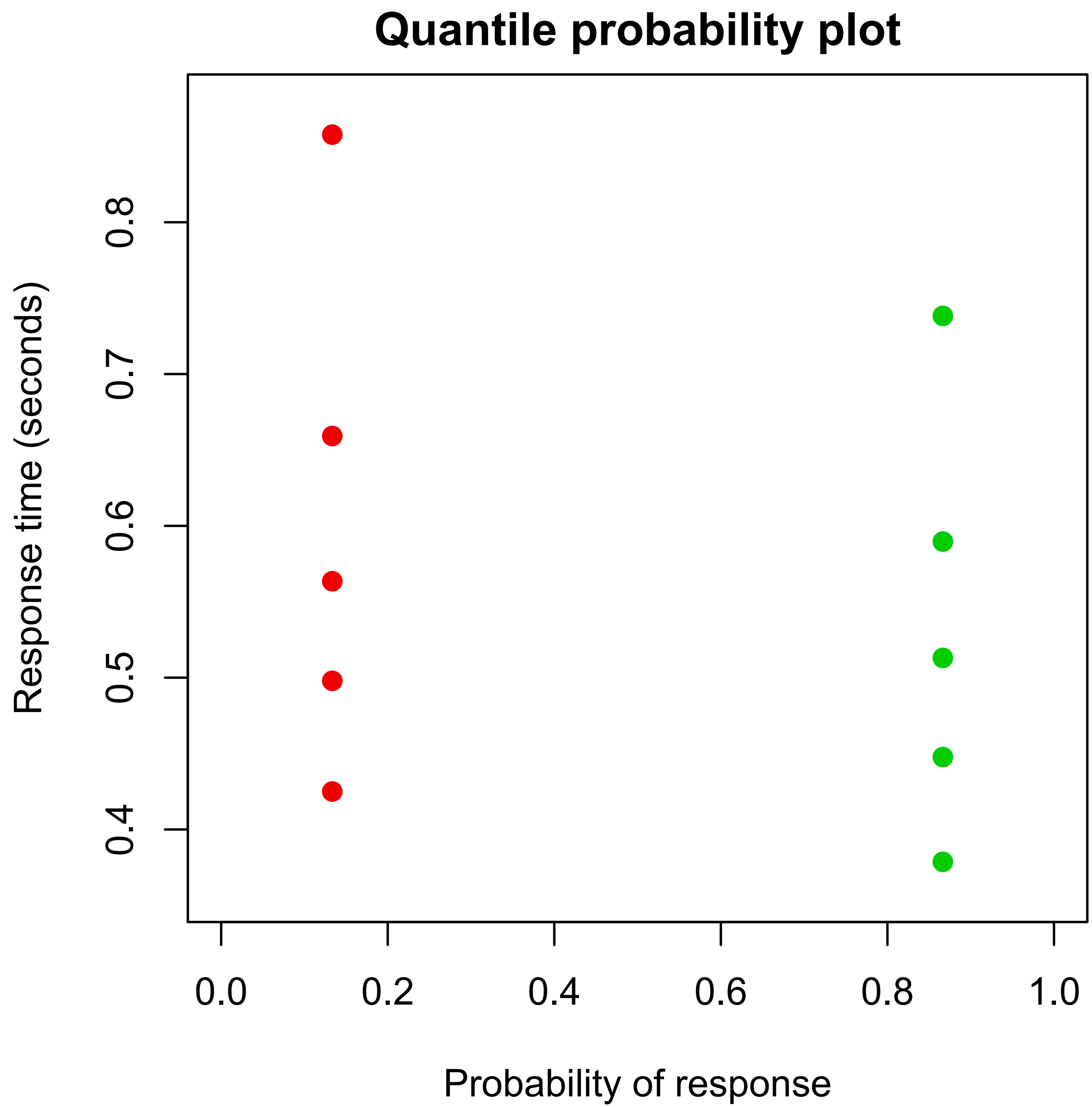
- But first, how do we plot the predictions?
  - Visualisation is key to evaluating models!


Mean RT
Response time (seconds)
0.56
0.58
0.60
0.62
Correct
Error


Histogram of RTs
Density
0.0
0.5
1.0
1.5
2.0
2.5
0.4
0.6
0.8
1.0
1.2
Response time (seconds)

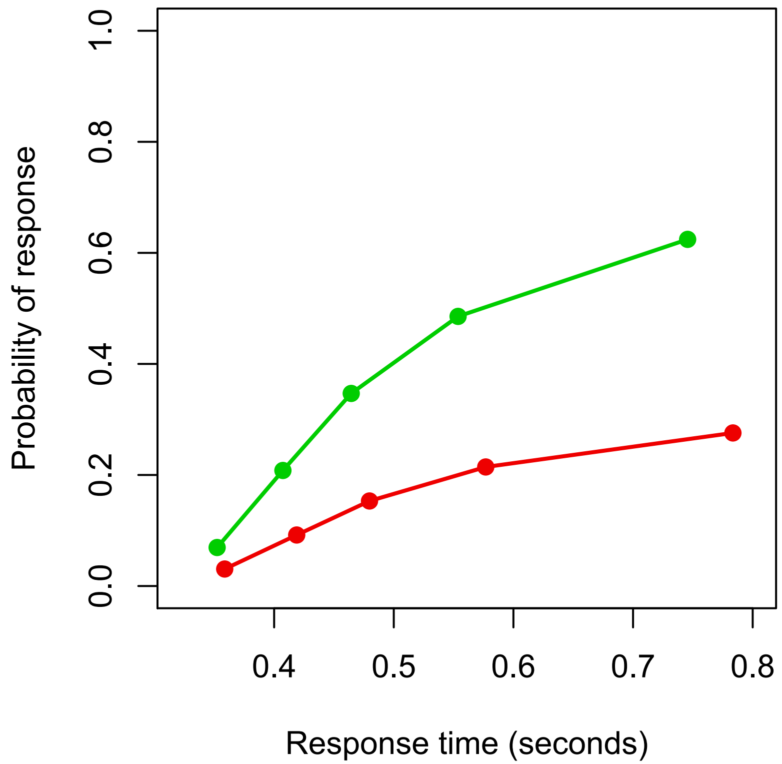


Defective CDF
Probability of response
Response time (seconds)
0.0
0.2
0.4
0.6
0.8
1.0
0.4
0.5
0.6
0.7
0.8

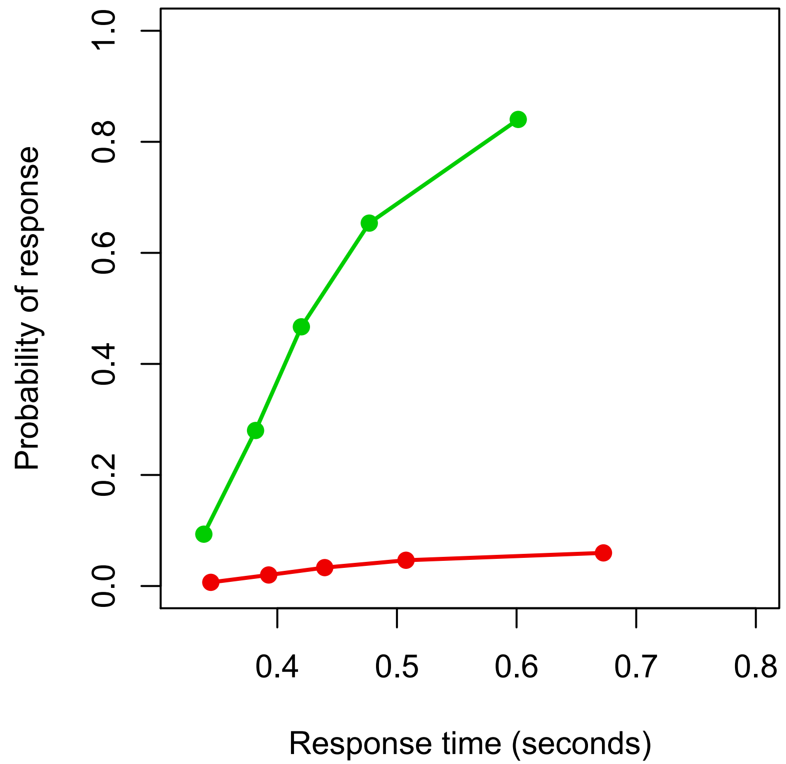

Quantile probability plot
Response time (seconds)
0.4
0.5
0.6
0.7
0.8
0.0
0.2
0.4
0.6
0.8
1.0
Probability of response

Drift rate

Lower drift rate



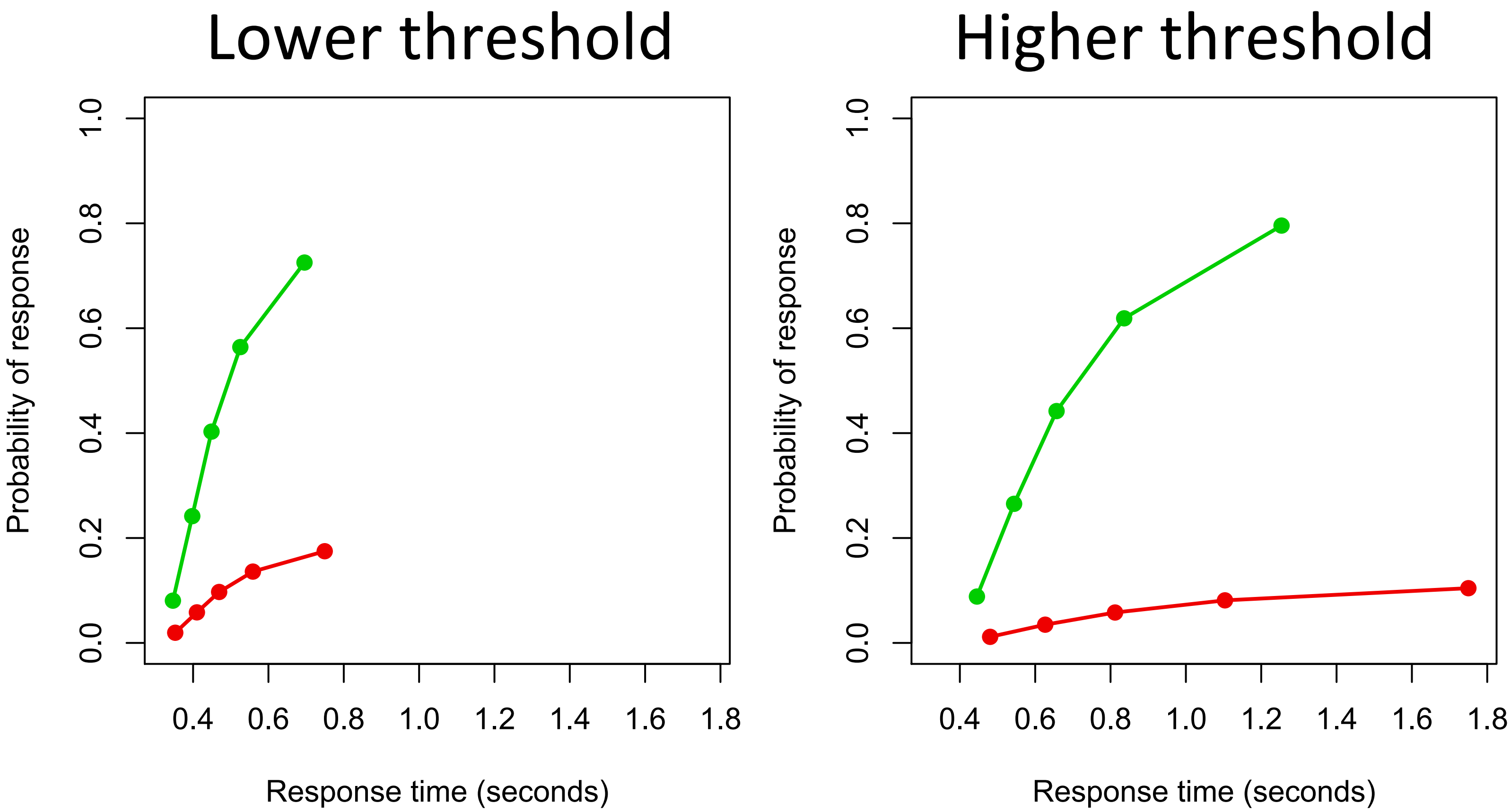
Higher drift rate



On average, increasing drift rate leads to:

- Faster responses
- More accurate responses

Response threshold

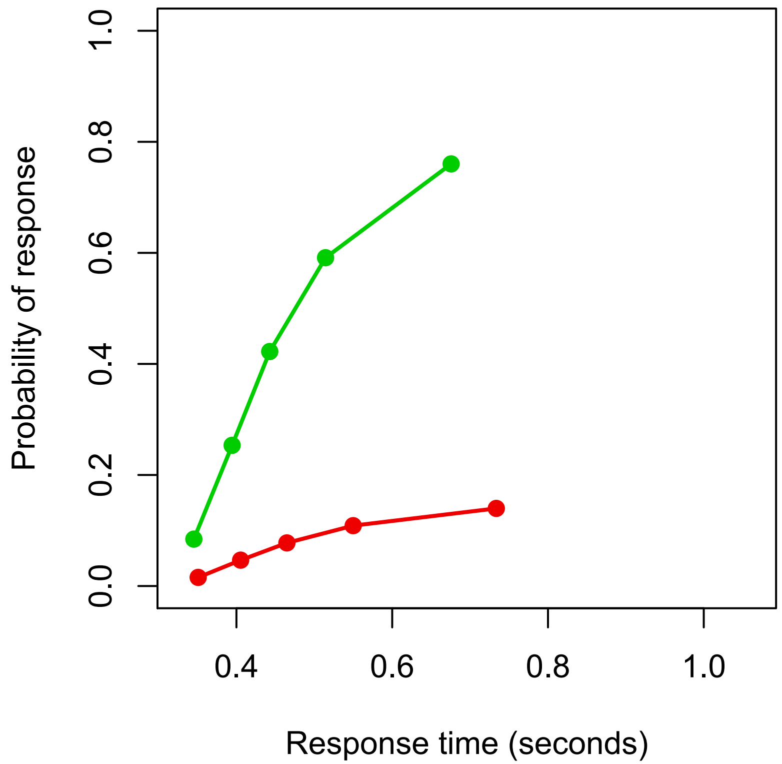


On average, increasing response threshold leads to:

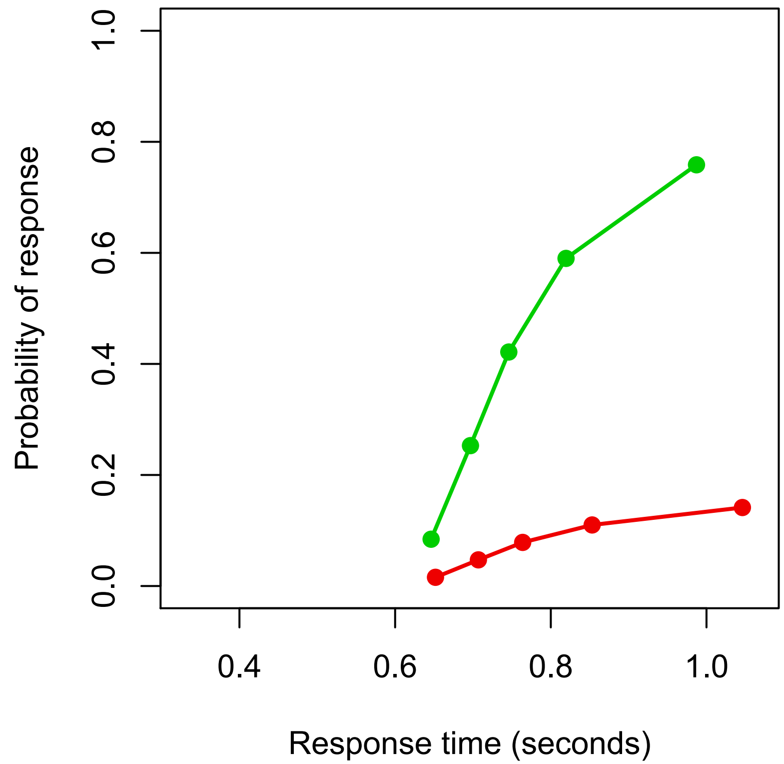
- Slower responses
- (slightly) more accurate responses

Non-decision time

Shorter non-decision time



Longer non-decision time



Increasing non-decision time leads to:

- Slower responses
- No change in accuracy

Practical

- Instructions and files: `osf.io/8zkb3`

- Open the file *practical-main.R* in Rstudio

- Exercise 1: Simulate data from LBA
  - Follow the code
  - Specify some parameter values
  - Simulate data
  - Summary statistics for simulated data
  - Plot simulated data in various ways

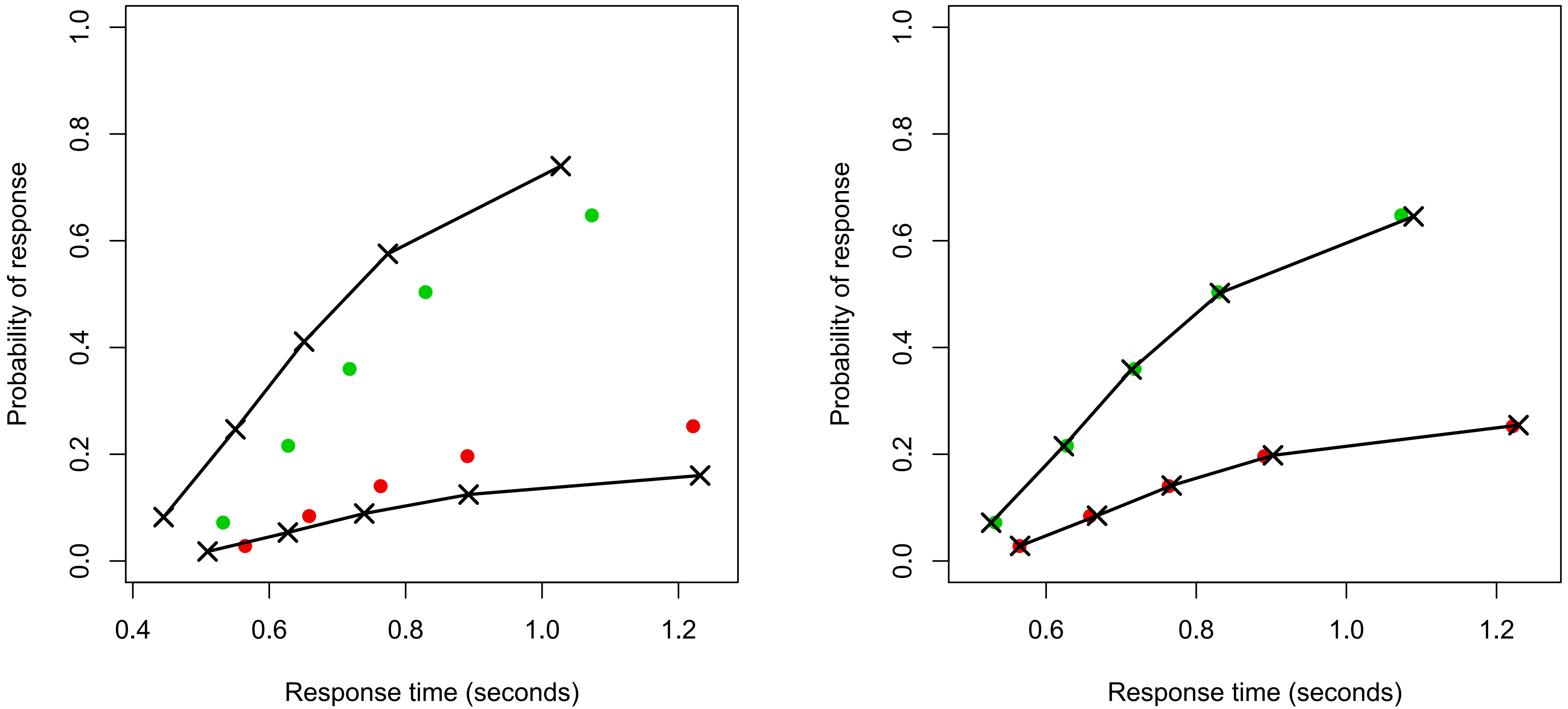
- Exercise 2: Simulate data from DDM
  - Same idea, but for DDM

Practical

- Exercise 3: Simulate data from an experiment with two conditions
  – Similar to Exercises 1 and 2
  – Adjust model parameters using vector notation
    - e.g.: `…, vc=c(3,2), ve=c(0,1), t0=0.2`

Fitting models to data

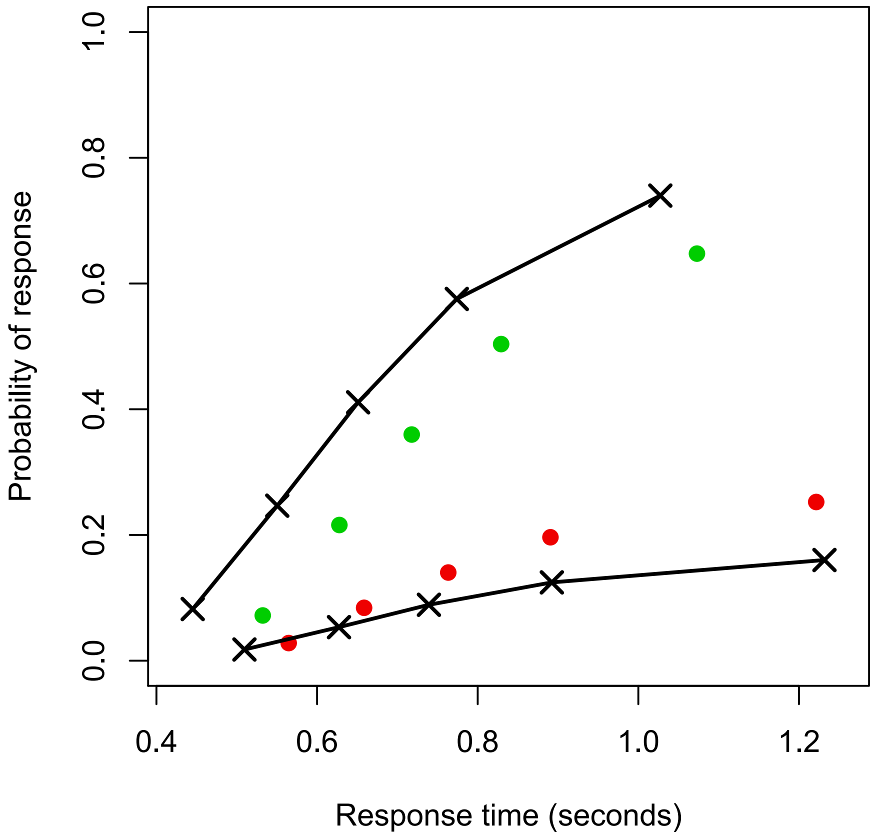
Aim: Identify the parameter values that minimise the ‘discrepancy’ between data and model predictions



Probability of response
0.0
0.2
0.4
0.6
0.8
1.0
0.4
0.6
0.8
1.0
1.2
Response time (seconds)
Probability of response
0.0
0.2
0.4
0.6
0.8
1.0
0.6
0.8
1.0
1.2
Response time (seconds)

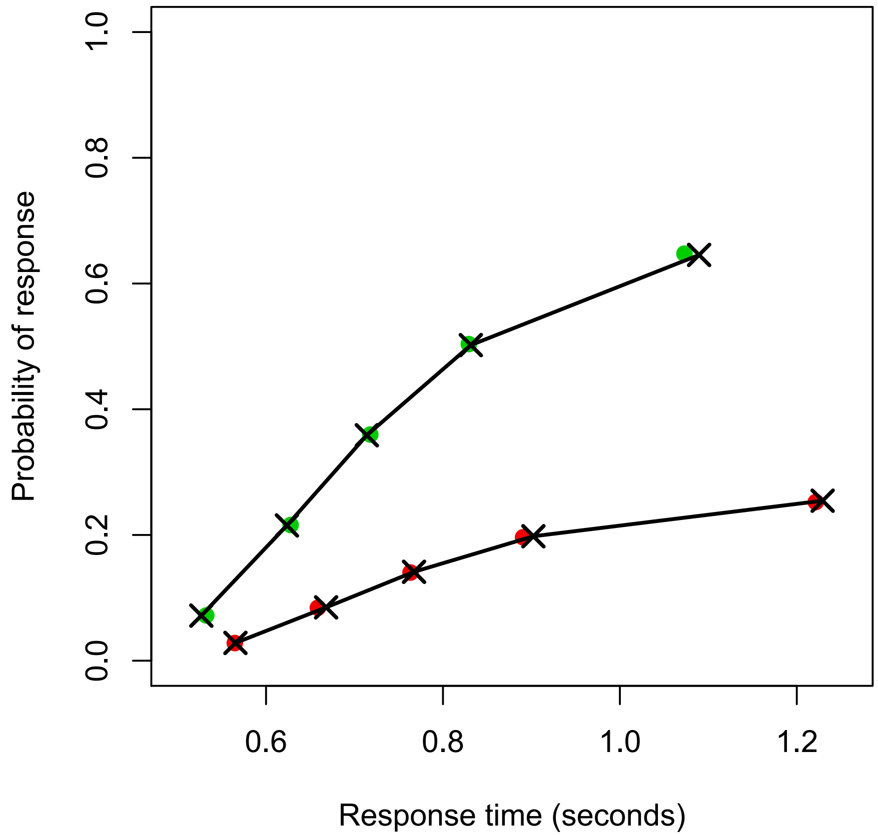
Large discrepancy

bad model fit



Small discrepancy

good model fit



Parameter estimation involves specification of

- A model
  - LBA, DDM, etc

- Parameter constraints
  - Some parameters may be freely estimated from data, others can be constrained or fixed to specific values
  - e.g., We could estimate a separate drift rate across two conditions but only a single response threshold

- Objective function
  - Must define how we measure the discrepancy or 'distance' between data and model predictions

Objective function

- Ideally we want likelihood (PDF) of each data point given model parameters
  - Likelihood of a response at a given time
  - Likelihood is a sufficient statistic
    - Carries all the information contained in the data
  - Available for LBA and DDM

Race equation (LBA)

$$\mathrm{PDF}_i(t) = f_i(t) \prod_{j \neq i} \big(1 - F_j(t)\big)$$

Race equation (LBA)

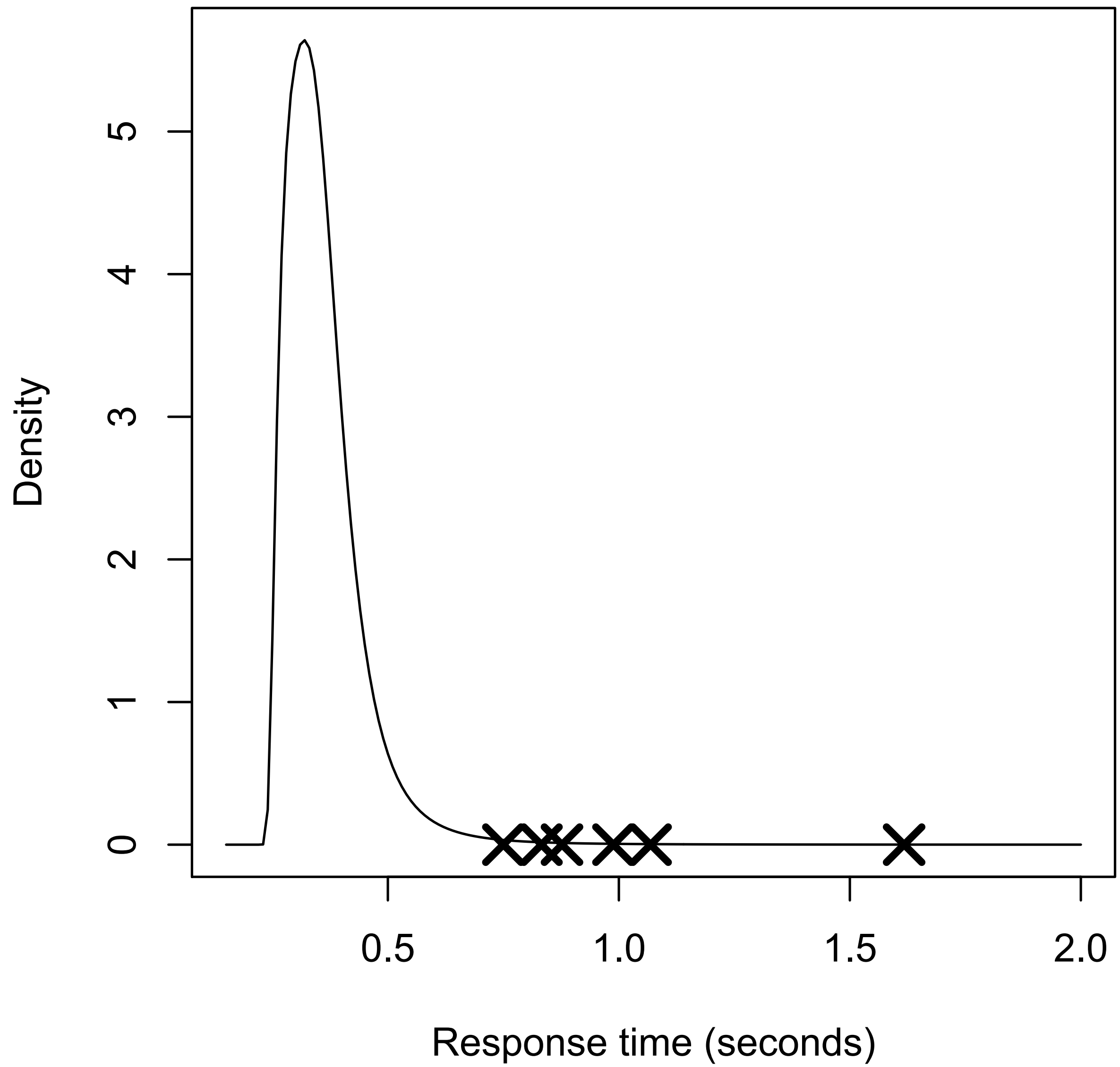
$$\mathrm{PDF}_i(t) = f_i(t) \prod_{j \neq i} \left(1 - F_j(t)\right)$$

$$f_i(t) = \frac{1}{A}\left[-v_i \Phi\left(\frac{b - A - tv_i}{ts}\right) + s\phi\left(\frac{b - A - tv_i}{ts}\right) + v_i \Phi\left(\frac{b - tv_i}{ts}\right) - s\phi\left(\frac{b - tv_i}{ts}\right)\right]$$

$$F_i(t) = 1 + \frac{b - A - tv_i}{A}\Phi\left(\frac{b - A - tv_i}{ts}\right) - \frac{b - tv_i}{A}\Phi\left(\frac{b - tv_i}{ts}\right) + \frac{ts}{A}\phi\left(\frac{b - A - tv_i}{ts}\right) - \frac{ts}{A}\phi\left(\frac{b - tv_i}{ts}\right)$$

Maximum likelihood estimation

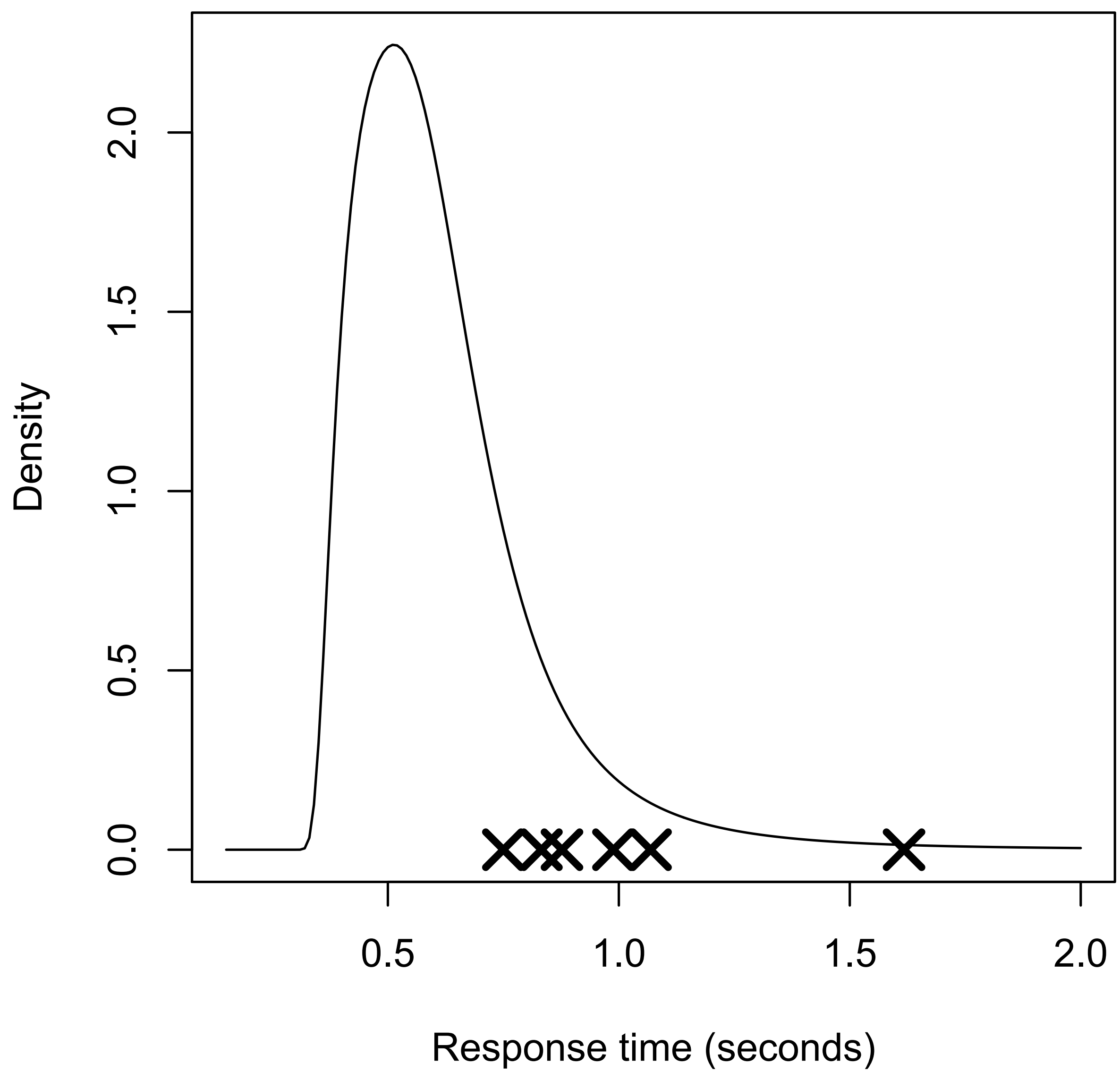
- Calculate likelihood of the data given a set of model parameters
  - *Likelihood* is the product of the density of each data point under the model PDF
  - *Log-likelihood* is the sum of the log of the density of each data point under the model PDF



Density
0
1
2
3
4
5
0.5
1.0
1.5
2.0
Response time (seconds)

Maximum likelihood estimation

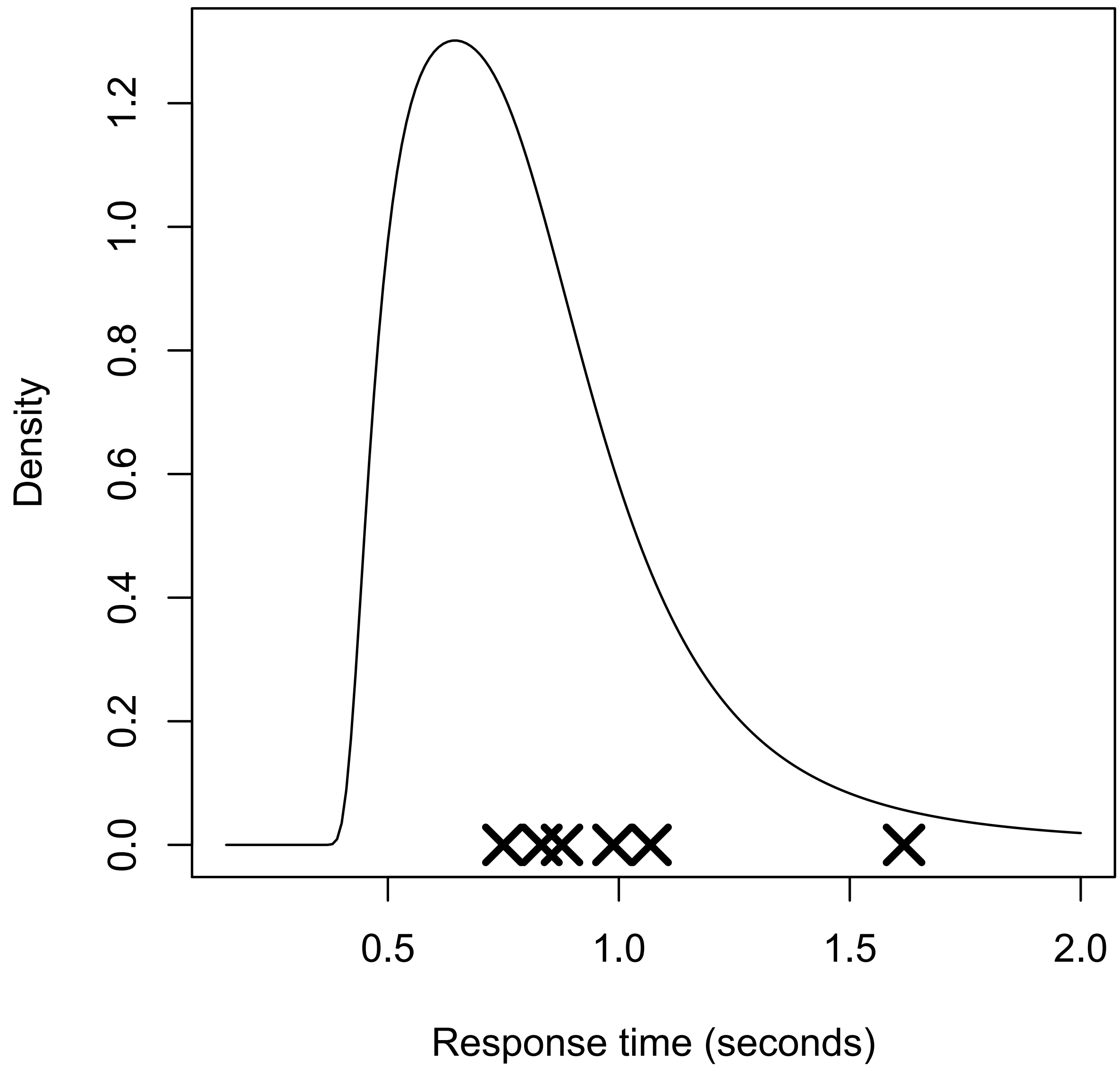
- Calculate likelihood of the data given a set of model parameters
- Adjust to a new set of parameters and test whether the likelihood increases



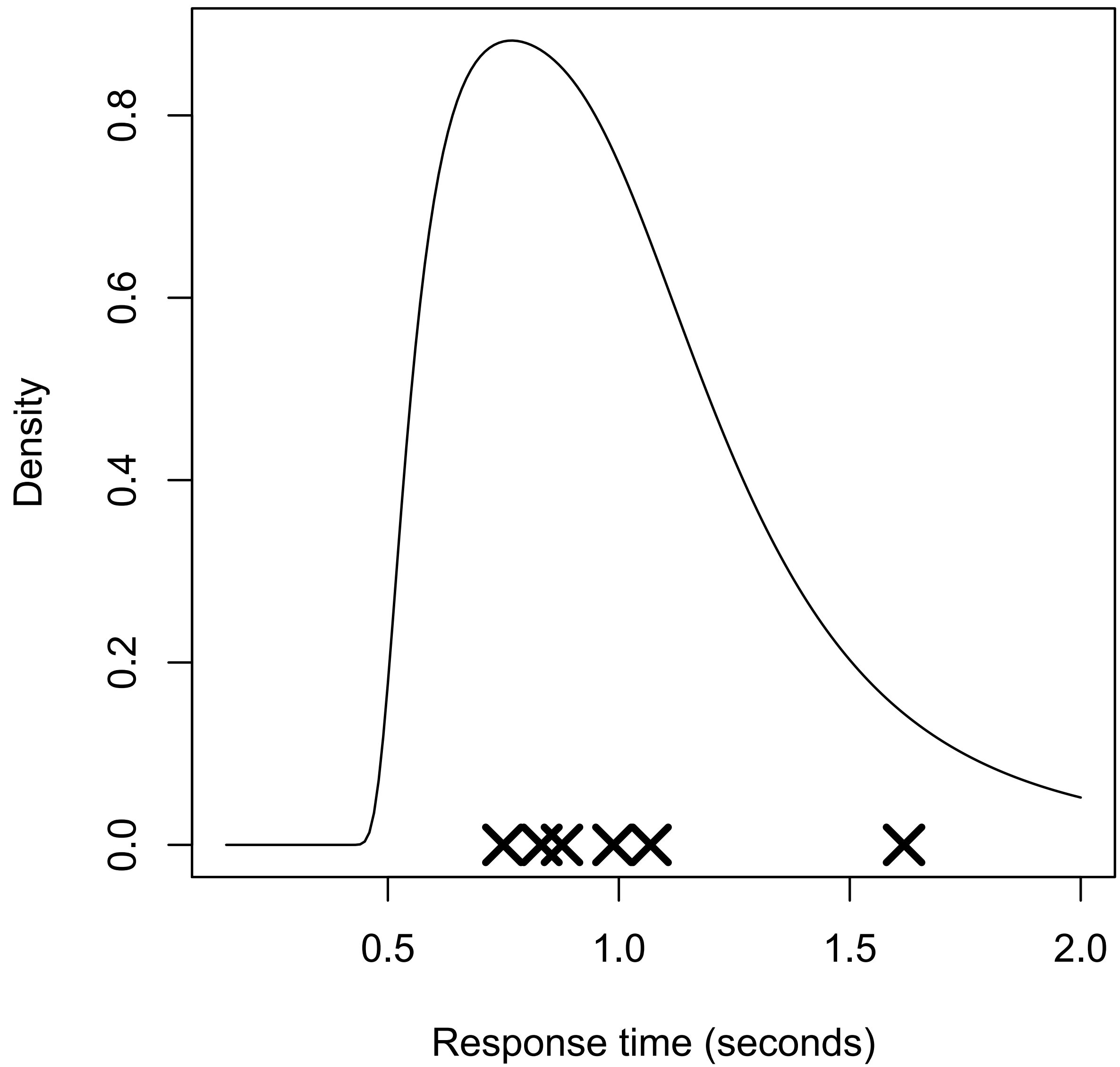
Density
0.0
0.5
1.0
1.5
2.0
0.5
1.0
1.5
2.0
Response time (seconds)

Maximum likelihood estimation

- Calculate likelihood of the data given a set of model parameters
- Use a particular rule to propose a new set of candidate parameters and test whether the likelihood increases
  - If it improves, accept the new parameters
- Repeat



Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0.5
1.0
1.5
2.0
Response time (seconds)



0.0
0.2
0.4
0.6
0.8
Density
0.5
1.0
1.5
2.0
Response time (seconds)

Practical

- Fitting exercise 1: Simulate data from LBA for a single experimental condition
  - Specify some parameter values
  - Simulate data
  - Specify maximum number of iterations for optimisation algorithm
    - Number of times candidate parameters are proposed
  - Optimise parameters
    - Note: Observe the value of the likelihood decrease over iterations
  - Evaluate fit to data
    - With summary statistics and plots

Parameter optimisation

- The rule/s for proposing and updating candidate sets of parameters differs across optimisation algorithms
  - There are *many* types of optimisation algorithms
  - Some are more efficient than others (faster computation time)

- We're using differential evolution
  - Simultaneously assesses many candidate sets of parameters
  - Easily implemented through the R library *DEoptim*

Optimisation isn't magic…

- Optimisation algorithms can get stuck in local minima
  - Climbing a hill while blindfolded
  - Always check results by re-running the optimisation algorithm using different start points
- Running for too few iterations can give a suboptimal fit
- Just because an optimisation algorithm returns a best fitting set of parameters, doesn't mean those parameters provide a good fit to data
  - ALWAYS check goodness of fit to data ➔ Draw figures
- No amount of parameter optimisation will fix a bad model

How do we select between models?

- Visual examination
- Model selection indices
  - Aim: Find the simplest model that provides a good fit to data
  - Only informative for comparisons across models fit to a particular data set
    - Not across different data sets
  - Akaike information criterion (AIC)
    - AIC = (-2 * log-likelihood) +
      (2 * number of parameters)
  - Bayesian information criterion (BIC)
    - BIC = (-2 * log-likelihood) +
      (log(number of data points) * number of parameters)

Practical

- Fitting exercise 2: Fit two models to a mystery data set to determine which model provides the best account of the data

Different ways to model decisions

- Mathematical models of decision making (e.g., LBA, DDM)
  - Pros
    - Tractable, generally good parameter estimation properties
    - Software programs exist to implement the models
  - Cons
    - Less biological plausibility

Different ways to model decisions

- Mathematical models of decision making (e.g., LBA, DDM)
  - Pros
    - Tractable, generally good parameter estimation properties
    - Software programs exist to implement the models
  - Cons
    - Less biological plausibility
- Simulation-based models of decision making (e.g., LCA)
  - Pros
    - Easy to get started on
    - Easily implement all kinds of biologically plausible ideas
  - Cons
    - Less tractable ➔ parameter estimation is difficult
      - No likelihood function - must use alternative statistics that vary in robustness and efficiency (if fit to data at all)
    - Hard to answer questions like: given this data set, what (combination of) parameters provide the best account of the effect?

Which software to use?

- Packaged software
  - DMAT, HDDM, EZ-diffusion
    - Easy to use, but
    - Black-box problem
    - Can have limited flexibility
- Customisable software
  - Matlab, Python, R
    - Flexible and allow users to develop libraries
    - Can have a steep learning curve
- We used R with a custom library that implements the math behind two common decision making models
  - Linear ballistic accumulator (LBA)
  - Diffusion decision model (DDM)

Conclusions

- Quantitative models provide a detailed account of cognition
  - Fitting models to data involves many steps
  - Today we've explored some of those steps at a conceptual level
    - The technical details were hidden. If you're interested, look at the file: *behind-the-scenes.R*