Quantitative Methods for Cognitive Scientists

Modeling Primer: Sequential Sampling Models

Emre Neftci

Department of Cognitive Sciences, UC Irvine,

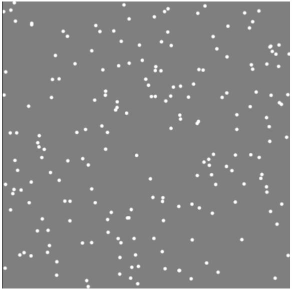
April 10, 2019

Emre Neftci

eneftci@uci.edu

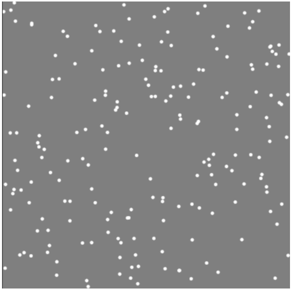
https://canvas.eee.uci.edu/courses/16991

Reaction time data provides a rich window into human cognition



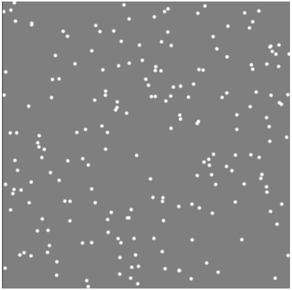
- Are the majority of the dots moving left or right? Respond as quickly as possible

Reaction time data provides a rich window into human cognition



- Are the majority of the dots moving left or right? Respond as quickly as possible
- Response data classified as correct vs. incorrect
- Each class is characterize by a distribution of Reaction Times (RT)

Reaction time data provides a rich window into human cognition



- Are the majority of the dots moving left or right? Respond as quickly as possible
- Response data classified as correct vs. incorrect
- Each class is characterize by a distribution of Reaction Times (RT)
- A perfect model of human cognition should describe response accuracy and latency and the relationship between the two.

- When a stimulus is presented, not all information is immediately available to the participant.
- We *gradually* build up evidence
- How would you do model this?

Modeling Assumptions

- Assume a model that samples evidence in discrete time windows

Modeling Assumptions

- Assume a model that samples evidence in discrete time windows
- Each sampled number represents a nudge toward one decision or another.

Modeling Assumptions

- Assume a model that samples evidence in discrete time windows
- Each sampled number represents a nudge toward one decision or another.
- The magnitude of that nudge reflects how much *information* is acquired in that single sample.

Modeling Assumptions

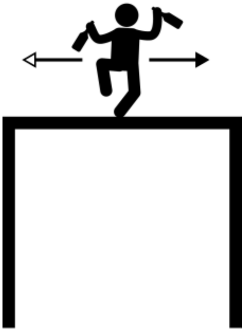
- Assume a model that samples evidence in discrete time windows
- Each sampled number represents a nudge toward one decision or another.
- The magnitude of that nudge reflects how much *information* is acquired in that single sample.
- The nudge would then be added to the sum of previous nudges, moving a decision variable towards the left or right.

Modeling Assumptions

- Assume a model that samples evidence in discrete time windows
- Each sampled number represents a nudge toward one decision or another.
- The magnitude of that nudge reflects how much *information* is acquired in that single sample.
- The nudge would then be added to the sum of previous nudges, moving a decision variable towards the left or right.
- A decision is made when the decision variable crosses a boundary

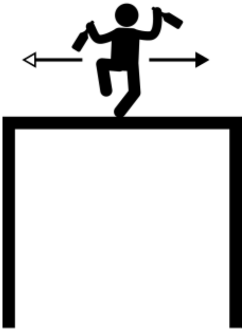
This defines a Random Walk model that stops when a boundary is reached

Every time step, a drunk person takes a left or right step on top of a cliff



We can ask two questions:

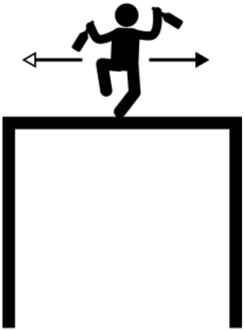
Every time step, a drunk person takes a left or right step on top of a cliff



We can ask two questions:

- Question 1: Which side of the cliff will he fall over from?

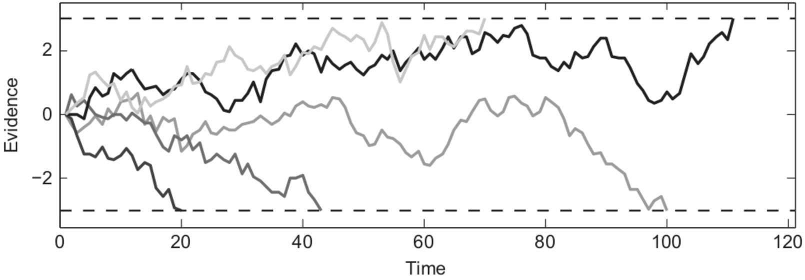
Every time step, a drunk person takes a left or right step on top of a cliff



We can ask two questions:

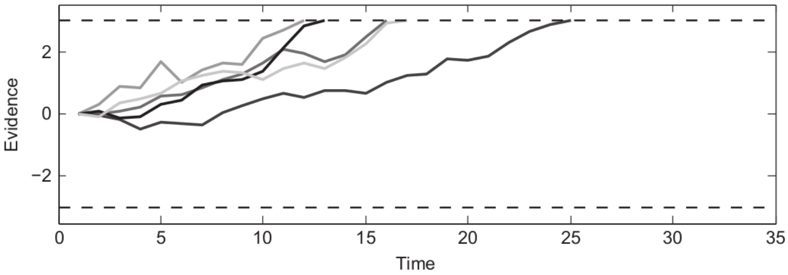
- Question 1: Which side of the cliff will he fall over from?
- Question 2: How long will he survive?

Five illustrative sampling paths when the stimulus is non-informative



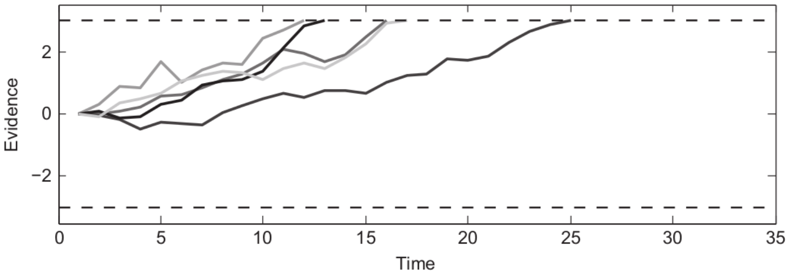
- Assume top boundary mean left response, bottom mean right
- What is the expected proportion of left vs. right responses
- What are the expected Reaction Time (RT) for left responses vs. right responses?

Five illustrative sampling paths when the stimulus is informative (i.e. there is nudging towards top)



- What is the expected proportion of left vs. right responses

Five illustrative sampling paths when the stimulus is informative (i.e. there is nudging towards top)



- What is the expected proportion of left vs. right responses
- What can we say about the RTs about incorrect responses?

Random walk model:

- Suppose the random walk process starts at X_0
- Suppose X_t is the nudge at time t. Assume X_t can be any positive or negative number, randomly drawn from some distribution
- The evidence after 2 steps is defined as: $Z_2 = X_0 + X_1 + X_2$

Mathematical Formulation of a Random Walk Model

Random walk model:

- Suppose the random walk process starts at X_0
- Suppose X_t is the nudge at time t. Assume X_t can be any positive or negative number, randomly drawn from some distribution
- The evidence after 2 steps is defined as: $Z_2 = X_0 + X_1 + X_2$
- More generally, after T steps $Z_T = \sum_{t=0}^{T} X_t$

... with boundaries at -3 and 3

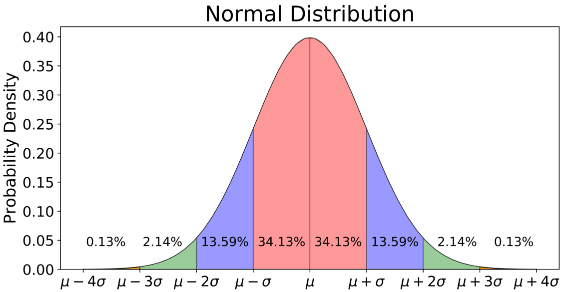
- If $Z_t > 3$ then the response is "left" and the simulation is stopped
- If $Z_t < -3$ then the response is "right" and the simulation is stopped
- The time t at which either left or right response is given is the response time

How can we represent "drift"? i.e. more evidence for left or right?

What does $X_0 \neq 0$ mean?

A Gaussian Random Walk

To make our random walk more specific, we assume that X_t are "drawn" from a Gaussian distribution with parameters μ_D (drift) and σ_D (error in the evidence).



The notation:

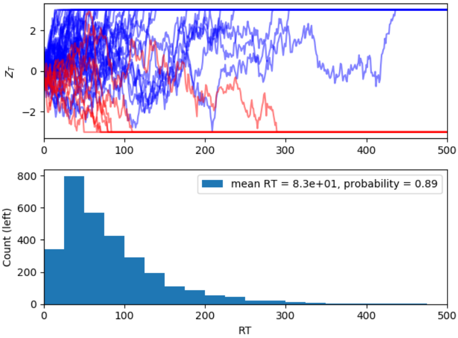
$$X_t \sim N(\mu_D, \sigma_D)$$

means X_t is distributed according to a normal (=Gaussian) distribution with mean μ_D and standard deviation σ_D.

(Note: later in this course we will see more about distributions)

Simulations of the Random Walk Model

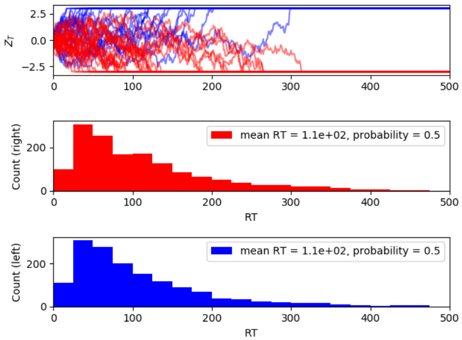
1000 trials of random walk model with
$X_t \sim N(\mu_D = 0, \sigma_D = 0.3)$, starting at zero ($X_0 = 0$)



- Assume top boundary = left, bottom boundary = right
- Bottom: Recorded RTs

Simulations of the Random Walk Model for zero drift

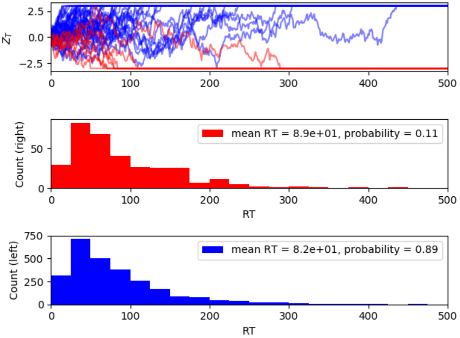
Drift is zero: $X_t \sim N(\mu_D = 0.0, \sigma = 0.025)$, starting at zero ($X_0 = 0$)



- Middle: Recorded RTs for right decisions
- Bottom: Recorded RTs for left decisions
- (remember right = bottom = incorrect, left = top = correct)

Simulations of the Random Walk Model for Positive Drift

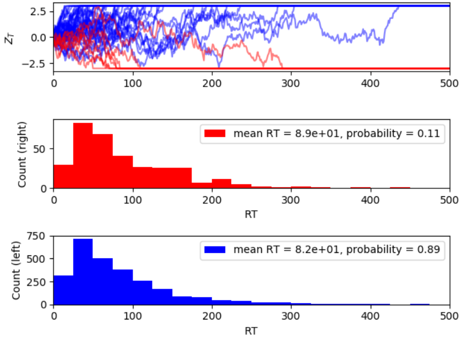
Drift is positive: $X_t \sim N(\mu_D = 0.03, \sigma = 0.025)$, starting at zero ($X_0 = 0$)



- More Left Responses

Simulations of the Random Walk Model for Positive Drift

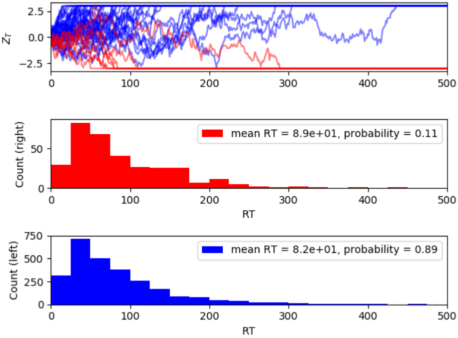
Drift is positive: $X_t \sim N(\mu_D = 0.03, \sigma = 0.025)$, starting at zero ($X_0 = 0$)



- More Left Responses
- Drift does not change reaction times! Why?

Simulations of the Random Walk Model for Positive Drift

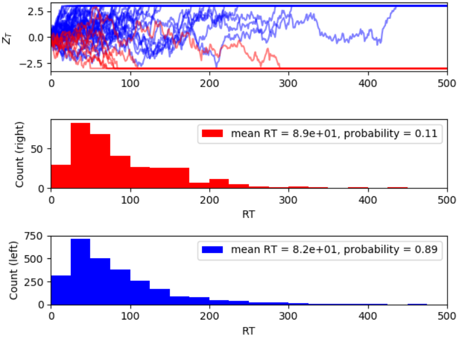
Drift is positive: $X_t \sim N(\mu_D = 0.03, \sigma = 0.025)$, starting at zero $(X_0 = 0)$



- More Left Responses
- Drift does not change reaction times! Why?
- The only errors the model can produce are those that occur as quickly as a correct response

Simulations of the Random Walk Model for Positive Drift

Drift is positive: $X_t \sim N(\mu_D = 0.03, \sigma = 0.025)$, starting at zero ($X_0 = 0$)

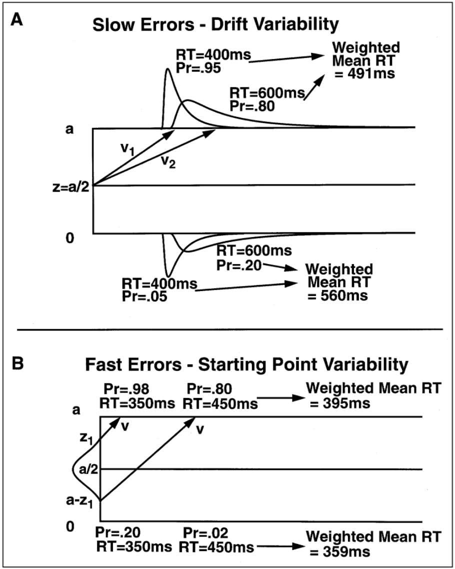


- More Left Responses
- Drift does not change reaction times! Why?
- The only errors the model can produce are those that occur as quickly as a correct response
- However, people's errors are often faster than correct responses

Trial-to-Trial Variability in the Random-Walk Model

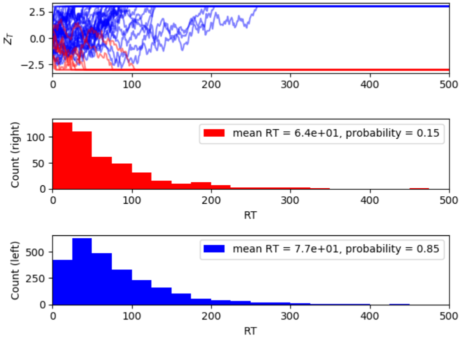
- Variability so far was due to the sampling process.
- We can introduce a different type of variability, due to differences from one trial to another
- In out model, trial-to-trial variability refers to changes in the values of parameters between different simulated trials.
- Let's assume randomness in the starting point (X_0) and the drift (μ_D)

Trial-to-trial Variability in the Random Walk Model



RatCliff and Rouder, Modeling response times for two-choice decisions, 1998

Trial-to-trial variability in the starting points
$X_0 \sim N(\mu_0 = 0, \sigma_0 = 1.0)$

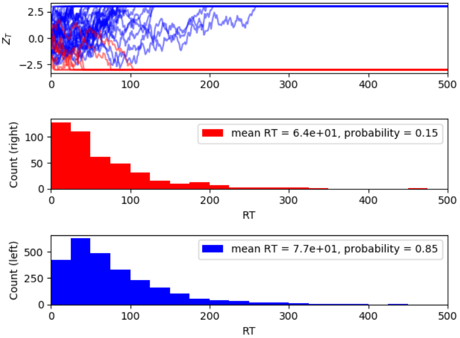


- Incorrect responses are faster

Simulations of the Random Walk Model with Trial-to-Trial Variability

Trial-to-trial variability in the starting points

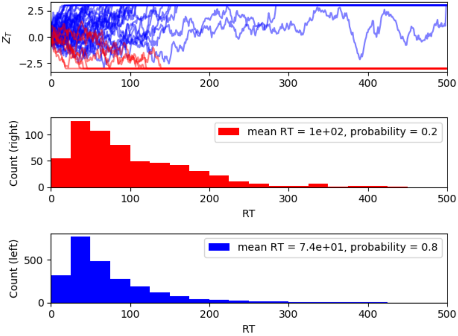
$X_0 \sim N(\mu_0 = 0, \sigma_0 = 1.0)$



- Incorrect responses are faster
- Interpretation: when errors arise, they are likely associated with a starting point close to the incorrect boundary and hence they are quick

Simulations of the Random Walk Model with Trial-to-Trial Variability

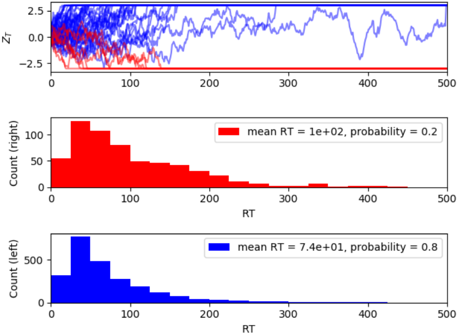
Trial-to-trial variability in drift ($\mu_D \sim N(0.00, .025)$)



- Incorrect responses are slower

Simulations of the Random Walk Model with Trial-to-Trial Variability

Trial-to-trial variability in drift ($\mu_D \sim N(0.00, .025)$)



- Incorrect responses are slower
- Interpretation: Incorrect responses can be due to a small drift towards the incorrect decision

The Random Walk model with trial-to-trial variability has the following "free" parameters:

- X_0: Average Starting point (for trial-to-trial variability)
- σ_0: Standard deviation of the starting point
- μ_D: Drift (for evidence in stimulus)
- σ: Sampling error accounting for noise in the evidence
- σ_D: Standard deviation of the Drift (for trial-to-trial variability)

Note that other parameters such as threshold, number of trials, variability in the stimulus are fixed by experimental design.

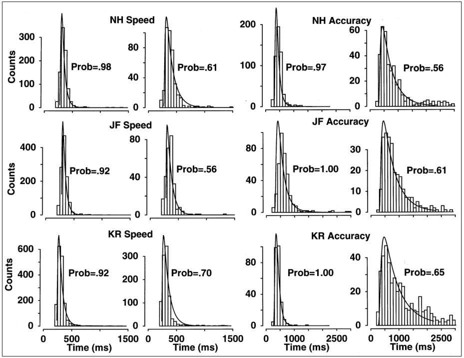
Response Times in Humans in Evidence Accumulation Tasks

PSYCHOLOGICAL SCIENCE

Research Article

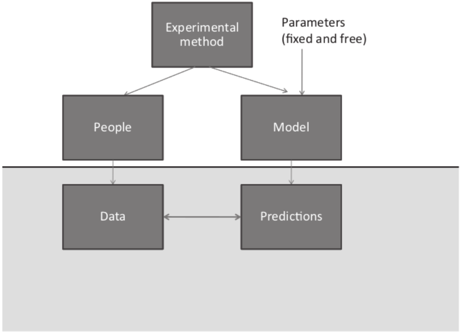
MODELING RESPONSE TIMES FOR TWO-CHOICE DECISIONS

Roger Ratcliff and Jeffrey N. Rouder
Northwestern University



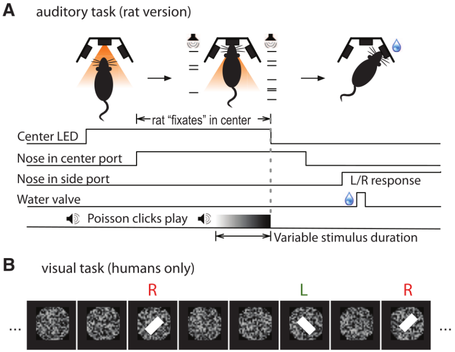
Ratcliff and Rouder, *Psychological science*, 1998

Summary: Connecting Model and Data



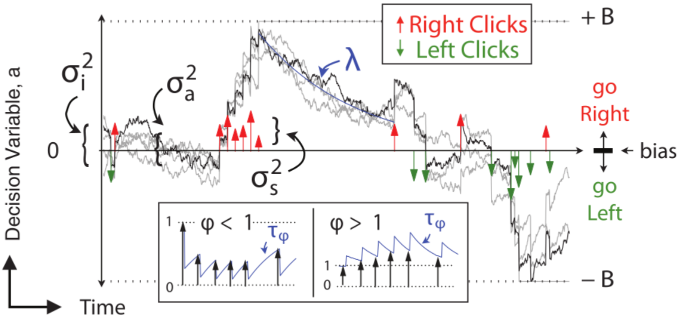
- In a research setting, one may record response times and decisions made over a large number of trials and individuals, using a fixed experimental setup.
- How can we estimate the free parameters for a group of people? for each individual?

Rats and Humans Can Optimally Accumulate Evidence for Decision-Making



Brunton, Botvinick, and Brody, *Science*, 2013

Trains of randomly timed clicks were played concurrently from left and right free-field speakers during the last portion of the fixation time.



Brunton, Botvinick, and Brody, *Science*, 2013

We reasoned that the precisely known pulse timing would enable detailed modeling of the subjects' choices on each individual trial, whereas its variability would allow exploration of the stimulus space and would thus provide statistical power.

Parameter Estimation Strategies: Function-based

Approach 1: Derive a function of the responses and the reaction times and fit the parameters of this function.

For example, the following probability function describes the distribution of incorrect reaction times:

$$P(t) = \frac{\pi}{a^2} \exp\left(-\mu_d X_0 - \frac{\mu_D^2 t}{2}\right) \sum_{k=1}^{\infty} k \exp\left(-\frac{k^2 \pi^2 t}{2a^2}\right) \sin(k\pi X_0 / a) \tag{1}$$

where a is the separation between upper and lower threshold (e.g. 6 in our case). This model assumes $\sigma = 1$.

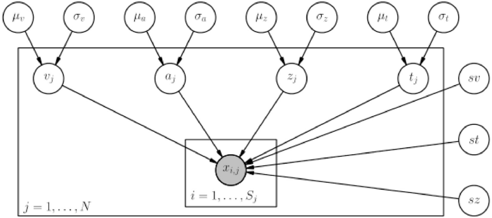
Navarro and Fuss, 2009

Pros & Cons:

- + Solutions are exact
- - It is hard, and often impossible to find such solutions. There are no analytical solution for the case with trial to trial variability.

Approach 2: Do many simulations with diffierent parameters. Choose the parameters of the simulation that had the best fit.

For example, using Hierarchical Bayesian Inference with Markov Chain Monte Carlo (we'll see the gist of Hierarchical Bayesian models later in this class)



Brunton, Botvinick, and Brody, *Science*, 2013

Pros & Cons:

+ Any model can be simulated
- Solutions are appromixate
- Computationally expensive

Further reading on Sequential Sampling Models:

- Farrell and Lewandowsky,, 2018, Chapter 2
- Brunton, Botvinick, and Brody, *Science*, 2013
- Wiecki, Sofer, and Frank, *Frontiers in neuroinformatics*, 2013
- Ratcliff and Rouder, *Psychological science*, 1998