Computational models of cognitive tasks

Benjamín Garzón

PhD course 3045, May 2018

Generative model

$$p(Y, \theta) = p(\theta)p(Y|\theta)$$

- Y observed variables (data)
  - Choices
  - Response times
- θ hidden, latent variables of interest (parameters)
- Designed by the analyst based on their assumptions

Generative model

In our case:

- X predictors (e. g. information from stimuli)
  $\alpha = f(\theta, X)$ (function of the parameters and the predictors)
  $p(Y|\theta) = distribution(\alpha)$

Warnings

- Parameter names may change (literature, code)
- Some implementation decisions are technical (convenient to allow estimation of the model) and not theoretical

Computational models

- Delay discounting
- Two-alternative forced choice (Drift diffusion model)
- Two-armed bandit
- Probabilistic reversal learning

Delay discounting

- Preference for immediate over delayed rewards
- Smaller-but-sooner over larger-but-later
- Aka temporal discounting or inter-temporal choice
- (Relatively) stable trait, large inter-individual variability
- Related to impulsivity and self-control. Impulsive choice (impatience) vs impulsive action

Relevance

Relevant to psychiatry and behavioural economics

- Addiction
- ADHD
- Pathological gambling
- Tobacco, drug and alcohol use
- Risky sexual practices
- Large credit card debt
- Also: depression, mania, anorexia nervosa, schizophrenia...

(Story et al., 2016)

Reasons for discounting

- Uncertainty in the delivery of future rewards
- Volatility of the environment
- Opportunity cost

(Story et al., 2016)

Questionnaire

Do you prefer?

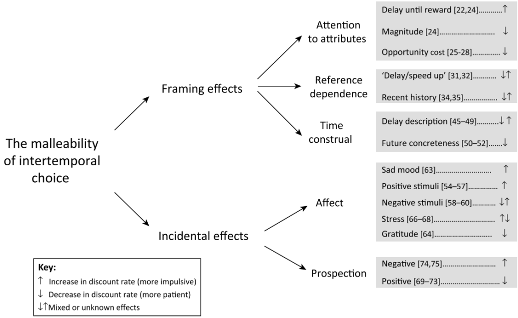
1. 1000 SEK now OR 2000 SEK in a month?
2. 1000 SEK now OR 1500 SEK in a month?
3. 1000 SEK now OR 2000 SEK in a year?
4. 10000 SEK in a week OR 20000 SEK in a year?

...

n. 10000 SEK now OR 20000 SEK in a year?

Limitations and manipulation

- Questionnaire decisions can be different from those faced in daily life
- Decisions are context-dependent
- Choices can be manipulated



(Lempert et al., 2016)

Subjective value

$$V(R) = f(d,\ R)$$

$f(d)$: decreasing function of the delay until R is received, e.g.:

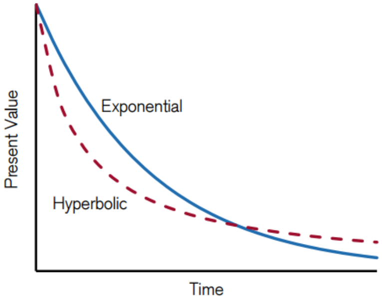
$$V(1000\ SEK) = f(1\ month, 1000\ SEK) = 900\ SEK$$

Exponential discounting

- Value declines exponentially with delay
- *k* is the *discount rate* (higher *k* = more impatient)
- Dynamic consistency:

  Discount equally with shorter and longer delays
  20 SEK in 1 week > 25 SEK in 2 weeks, 20 SEK in 9 weeks > 25 SEK in 10 weeks

$$V(R) = Re^{-kd}$$



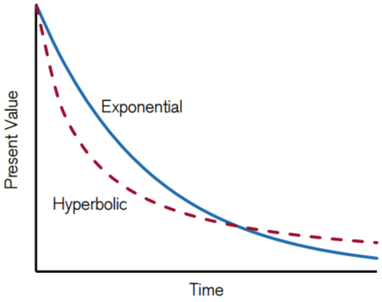
Hyperbolic discounting

- Value declines hyperbolically with delay
- Dynamic inconsistency:

  Discount more steeply with shorter delays and more shallowly with longer delays

  20 SEK in 1 week > 25 SEK in 2 weeks, but 25 SEK in 10 weeks > 20 SEK in 9 weeks

$$V(R) = \frac{R}{1 + kd}$$



Model

For trial t:

- Value V_{St}, reward R_{St} and delay d_{St} of sooner option
- Value V_{Lt}, reward R_{Lt} and delay d_{Lt} of later option

Value model:

$$V_{St} = \frac{R_{St}}{1 + kd_{St}} \qquad V_{Lt} = \frac{R_{Lt}}{1 + kd_{Lt}}$$

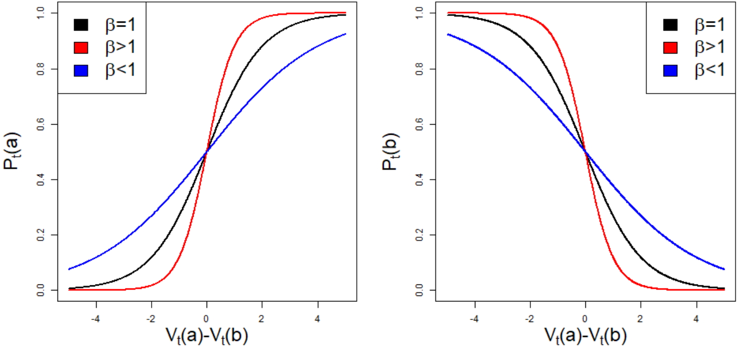
Action selection

Softmax function

$P_t(c)$: probability of choice $c \in a, b$

$$P_t(a) = \frac{e^{\beta V_t(a)}}{e^{\beta V_t(a)} + e^{\beta V_t(b)}} = \frac{1}{1 + e^{-\beta(V_t(a) - V_t(b))}}$$

$$P_t(b) = \frac{e^{\beta V_t(b)}}{e^{\beta V_t(a)} + e^{\beta V_t(b)}}$$



Model

Value model:

$$V_{St} = \frac{R_{St}}{1 + kd_t} \qquad\qquad V_{Lt} = \frac{R_{Lt}}{1 + kd_t}$$

Choice model:

$$c_t \sim bernoulli(softmax(\beta(V_{Lt} - V_{St})))$$

Built-in functions

```
Constants: e(), pi(),...
min(x), max(x)
exp(x), log(x), pow(x , y)
cos(x), sin(x), tan(x)
logit(y), inv_logit(x)
inv_Phi(y), Phi(x), Phi_approx(x)  //probit
```

Choice model implementation

Choice model:

$$c_t \sim bernoulli(softmax(\beta(V_{Lt} - V_{St})))$$

```
model {
 y ~ bernoulli(inv_logit(beta * (V_lt - V_st)));
 ...
}
model {
 y ~ bernoulli_logit(beta * (V_lt - V_st));
 ...
}
```

Data and parameters

Data

- Choices c_t
- Rewards R_{St}, R_{Lt}
- Delays d_{St}, d_{Lt}

Parameters

- Discount rate k
- Inverse temperature β
- (Expected values V_{St}, V_{Lt})

Drift diffusion model (DDM)

- Response-time data (relationship between choice accuracy and response time)
- Two-alternative forced choice tasks (e.g. left / right)
- Comparison process is not deterministic, choices not always optimal
- Typically perceptual tasks, but also value-based decision making
- Intuitive and biologically plausible
- Sequential sampling models (SSM)
- No error feedback

(Ratcliff et al., 1978)

Examples

Perceptual discrimination

- Random dot motion perceptual discrimination task

Numerosity

- Between 31 and 70 asterisks placed in random positions in a 10 x 10 array
- Choices: Small (<=50); large (>50)
- 30 blocks x 40 trials

Recognition memory

- 26 study-test blocks
- Study: 16 words
- Test: 16 studied + 16 new words

(Ratcliff et al., 2010)

Examples

Emotional flanker task

- Target affective words (positive/negative) with flanker words above and below (congruent, incongruent, neutral)
- Subjects judged if the target word was positively or negatively valenced
- 1 practice block x 20 trials + 4 blocks x 30 trials
- Total set of ffective words: 5 positive, 5 negative, 4 neutral

(Pe et al., 2013)

Examples

Food choices

- Choices between pairs of appetitive snack foods
- They had to eat the food that they chose in a randomly selected trial
- Items previously rated by the subjects
- High- and low-pressure conditions

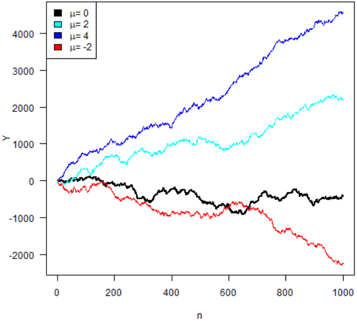
(Milosavljevic et al., 2013)

Model assumptions

- (Noisy) evidence about alternatives is accumulated over time
- Evidence = information from the environment or internal representations
- A choice is made when enough evidence has been gathered
- Non-decision time: not all response time is spent in accumulating evidence
- One of the responses may be more likely even before any evidence is observed
- Parameters are constant throughout the task

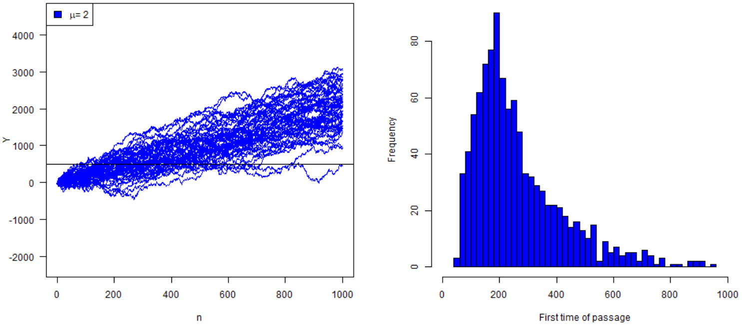
Evidence accumulation as a Wiener process

- $X_n \sim \mathcal{N}(\mu, \sigma^2)$, evidence at time step n
- $Y_0 = 0$; $Y_n = \sum_{k=1}^{n} X_k, n > 1$, total accumulated evidence at time step n



Evidence accumulation as a Wiener process

First passage time of the accumulation process



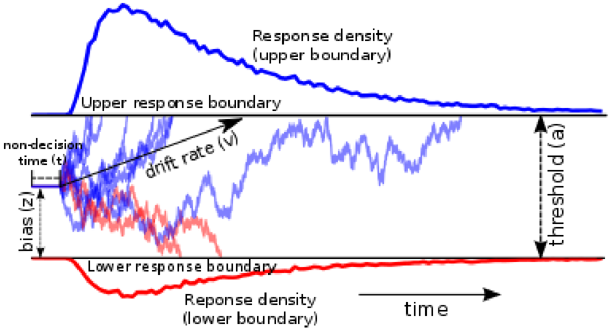
First passage time density function

$$p \sim W(a, v, z, t)$$

Choice model

Parameters

- Threshold a: Amount of evidence needed (response caution / speed-accuracy trade-off)
- Drift rate v: Rate of evidence accumulation (speed or efficiency of information processing / ability / task difficulty)
- Bias z: Information already present in favour of one of the choices ($0 < z < 1$, fraction of a)
- Non-decision time t: perception + motor execution + ...



(Wiecki et al., 2013)

Model predictions

DDM provides a simultaneous account of response time and accuracy

Noise can push the process toward the opposite direction, leading to an erroneous choice

Difficult decisions (small drift rate)

- Low accuracy
- Long response times

Easy decisions (large drift)

- High accuracy
- Short response times

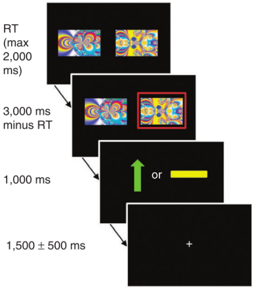
Effect of parameters

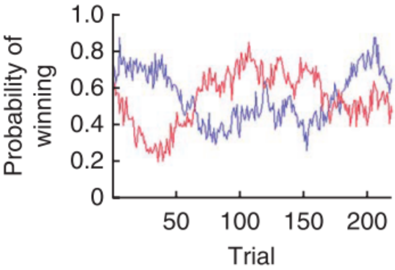
Model parameters affect differently RT and accuracy

- Increasing non-decision time increases RT
- Increasing drift rate decreases RT and increases accuracy
- Increasing the threshold increases both RT and accuracy

Two-armed bandit task

- Subjects choose between two options (bandits) on each trial, and receive a reward or punishment after the choice
- Subjects are instructed to maximize rewards
- The probability of obtaining a reward associated with each option varies on a trial-by-trial basis according to e.g. a Gaussian random walk

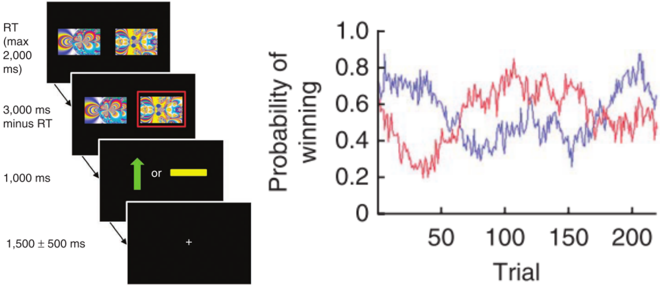




(Chowdhury et al., 2013)

Two-armed bandit task

- Assumption: to solve the task, subjects must keep an internal representation of the value of the stimuli, V_t
- We are interested in the learning rate α (how much subjects update their value representation based on rewards/punishments)
- Model can be extended in many ways



(Chowdhury et al., 2013)

Definitions

- c_t: choice (chosen stimulus) on trial t, $c_t \in \{a, b\}$
- r_t: reward on trial t, $r_t \in \{-1, 1\}$
- $V_t(c)$: value of choice $c \in \{a, b\}$ on trial t

Value model

Reward prediction error, Rescorla-Wagner rule:

$$\delta_t = r_t - V_t(c_t)$$

Value update (only chosen option is updated):

$$V_{t+1}(c_t) = V_t(c_t) + \alpha\delta_t$$

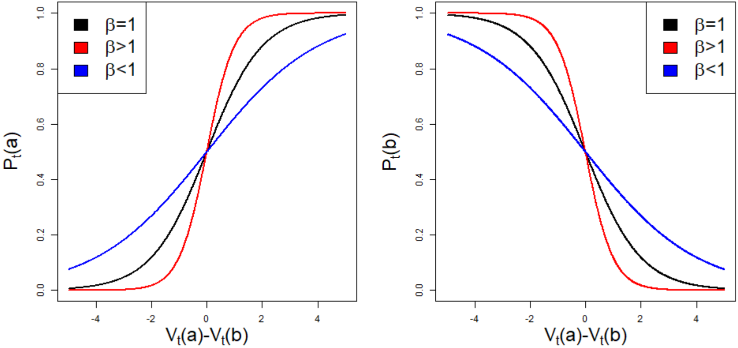
(Chowdhury et al., 2013)

Choice model

Softmax function

$P_t(c)$: probability of choice $c \in a, b$

$$P_t(a) = \frac{e^{\beta V_t(a)}}{e^{\beta V_t(a)} + e^{\beta V_t(b)}} \qquad P_t(b) = \frac{e^{\beta V_t(b)}}{e^{\beta V_t(a)} + e^{\beta V_t(b)}}$$



Data and parameters

Data

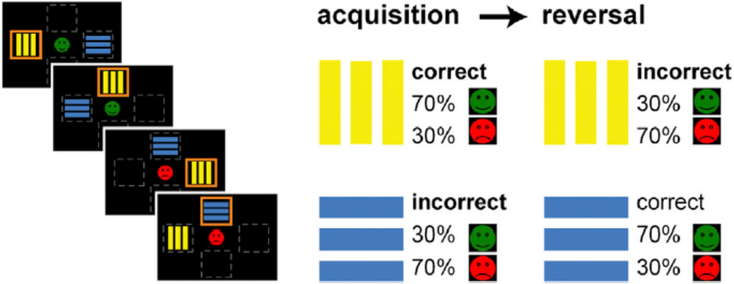
- Choices c_t
- Rewards r_t

Parameters

- Learning rate α
- Inverse temperature β
- (Expected values V_t)

Probabilistic reversal learning task

- Correct choices lead to monetary gains with probability p and losses with probability $q = 1 - p$ (e. g. $p = 0.7$ and $q = 0.3$)
- Incorrect choices lead to monetary gains with probability q and losses with probability p
- Subjects instructed to learn by trial and error to maximize reward
- Reward contingencies reverse at fixed points (e.g. 40 trials acquisition + 40 trials reversal) or after a number of consecutive correct choices
- Stimuli presented randomly at different spatial locations



(den Ouden et al., 2013)

Reversal learning

- Main interest is in adaptation to changes in stimulus-reward contingencies to study perseverative behaviour (Parkinson's disease, association with neurotransmitters)

(den Ouden et al., 2013, Cools et al., 2006)

Reward-punishment model

- Hypothesis: perseverative behaviour caused by reduced learning from punishment
- Separate learning rates for punishment α_p and reward α_r

Reward prediction error:

$$\delta_t = r_t - V_t(c_t)$$

Value update (only chosen option is updated):
If $r_t = 0$ (punished)

$$V_{t+1}(c_t) = V_t(c_t) + \alpha_p \delta_t$$

If $r_t = 1$ (rewarded)

$$V_{t+1}(c_t) = V_t(c_t) + \alpha_r \delta_t$$

Action selection: softmax function

(den Ouden et al., 2013)

Data and parameters

Data

- Choices c_t
- Rewards r_t

Parameters

- Learning rates α_r, α_p
- Inverse temperature β
- (Expected values V_t)

Experience-weighted attraction model

Hypothesis: perseveration on reversal is produced by an increasing reluctance to update the value of stimuli/choices every time they are chosen

Experience weight:

$$w_{t+1}(c_t) = w_t(c_t)\rho + 1$$

Value update:

$$V_{t+1}(c_t) = \frac{V_t(c_t)\varphi w_t(c_t) + r_t}{w_{t+1}(c_t)}$$

Action selection: softmax function

(Camerer et al., 1999)

Experience-weighted attraction model

Data

- Choices c_t
- Rewards r_t

Parameters

- Learning rate ϕ (= 1 - α)
- Experience decay factor ρ
- Inverse temperature β
- (Experience weights w_t)
- (Expected values $\overline{w_t}$)

Fitting models using hBayesDM

```
library(hBayesDM)
output = dd_hyperbolic("data/dd_exampleData.txt",
 niter=3000,
 nwarmup=1000,
 nchain=4,
 ncore=4)
```

References

- A. L. Odum. *Delay discounting, I'm a K, you're a K*, 2011
- G. W. Story, M. Moutoussis, R. J. Dolan. *A Computational Analysis of Aberrant Delay Discounting in Psychiatric Disorders*, 2016
- P. Samuelson. *A note on the measurement of utility*, 1937
- R Ratcliff. *A theory of memory retrieval*, 1978
- R. Ratcliff, A. Thapar, G. McKoon. *Individual differences, aging, and IQ in two-choice tasks*, 2010
- T. V. Wiecki, J. Poland, M. J. Frank. *Model-Based Cognitive Neuroscience Approaches to Computational Psychiatry: Clustering and Classification*, 2015
- M. L. Pe, J. Vandekerckhove, P. Kuppens. *A diffusion model account of the relationship between the emotional flanker task and rumination and depression*, 2013
- A. Diederich, P. Oswald. *Sequential sampling model for multiattribute choice alternatives with random attention time and processing order*, 2014
- U. Boehm, G. E. Hawkins, S. Brown, H. van Rijn, E. Wagenmakers. *Of monkeys and men: Impatience in perceptual decision-making*, 2016

References

- R. Chowdhury, M. Guitart-Masip, C. Lambert, P. Dayan, Q. Huys, E. Düzel, R. Dolan. *Dopamine restores reward prediction errors in old age*, 2013
- H. den Ouden, N. Daw, G. Fernández, J. A. Elshout, M. Rijpkema, M. Hoogman, B. Franke, R. Cools. *Dissociable effects of dopamine and serotonin on reversal learning*, 2013
- J. Gläscher, A. Hampton, J. O'Doherty. *Determining a Role for Ventromedial Prefrontal Cortex in Encoding Action-Based Value Signals During Reward-Related Decision Making*, 2009
- Stan Development Team. *Stan Modeling Language User's Guide and Reference Manual 2.16.0*, 2017
- http://mc-stan.org/