Bayesian Estimation of a Cognitive Model for Human Decisions and Associated Response Times

Henrik Singmann (Henrik.Singmann@warwick.ac.uk)

http://singmann.org/

Background

Many real-life decisions can be described as requiring a balance between the time taken for the decision and the quality of the decision. Consider for example the task of buying a new electronic device such as a smartphone. You could simply make a quick and dirty decision by visiting your preferred retailer and buying the first device in your price range. On the other end of the spectrum you can invest an almost unlimited amount of time before deciding on a specific device: you can visit several retailers and compare prices and features, you can read reviews, visit an online community to read about user experiences, contact the customer support, and so forth. Often our decisions are somewhere in the middle of this continuum; we do not take the first option we encounter, but we also make a decision within finite time.

The observation that decision makers can trade decision speed versus decision accuracy can also be observed on a smaller time scale. Consider for example an operator of an airport baggage scanner. When there are a lot of passengers they might be tempted to inspect each piece of baggage less carefully thus spending less time on them even if this is not necessary in line with the job affordances. Similar behaviours can be observed in many laboratory tasks used in psychological research. One way to investigate this is by manipulating the speed-accuracy trade-off via instructions (see Figure 1).

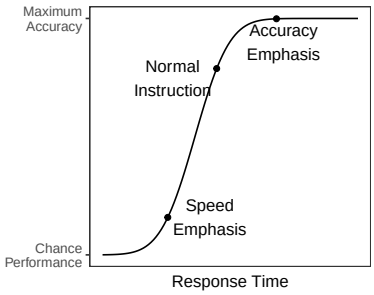

Figure 1: Idealized speed-accuracy operating characteristic function

The main formal approach for describing such speed-accuracy trade-offs is via evidence accumulation (or sequential sampling) models (Donkin and Brown, 2018). The general idea of this model class is the assumption that on encountering a stimulus, the decision maker accumulates noisy evidence using one or several evidence counters. Once an evidence counter reaches a pre-specified response criterion, the associated response is invariably given. By changing the position of the response criterion from liberal to strict the decision maker can trade-off speed versus accuracy. The fact that evidence accumulation is noisy provides the system with the ability to make errors.

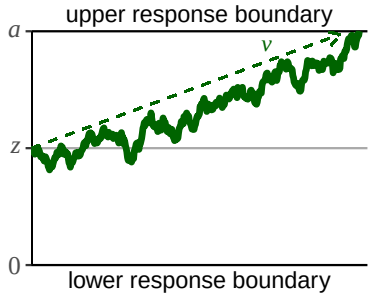

Figure 2: Basic diffusion model with three parameters: a = boundary separation, z = start point, v = drift rate.

The diffusion decision model (DDM, Figure 2) (Ratcliff and McKoon, 2008; Ratcliff et al., 2016; Ratcliff, 1978) is the most prominent variant of an evidence accumulation model. It assumes one single accumulator that accumulates evidence for one of two response options: positive evidence provides evidence for the response associated with the upper response boundary and negative evidence provides evidence for the response associated with the lower response

boundary. Evidence accumulation starts at start point z and once it reaches one of the two response boundaries, either 0 or a, the associated response is given. Furthermore, the DDM assumes that evidence is accumulated in a continuous fashion (compared to in discrete steps) with v representing the average drift towards one of the two options (e.g., if $v > 0$ average evidence drifts towards the upper response).

Project Outline

Fitting the DDM to real data from human decision makers is not always easy. One limitation is that when estimated in a Bayesian statistical framework using Markov chain Monte Carlo (MCMC) methods many earlier general purpose samplers appear to be inefficient. For example, posterior distributions from DDMs estimated using WinBUGS generally show a high-degree of autocorrelation requiring large amount of samples which can be prohibitively slow (Vandekerckhove et al., 2011). Such problems are usually exacerbated when specifying models in a hierarchical fashion.

Because of the problems with these earlier general purpose samples, proponents of DDMs have begun to develop custom samplers (Gunawan et al., 2020; Turner et al., 2013; Turner and Van Zandt, 2012; Heathcote et al., 2019). One of the downside of this approach is that in contrast to general purpose samplers available in probabilistic programming languages, these custom samplers are generally not validated or extensively tested.

The goal of this project is to assess the performance of the MCMC samplers available in the probabilistic programming language `NIMBLE` (de Valpine et al., 2017) for estimating the DDM to real data from psychological experiments. `NIMBLE` is a relatively recent probabilistic programming language that is fast (i.e., models are compiled in `C++`), convenient (i.e., provides an interface to `R`), flexible (i.e., provides a number of different MCMC samplers), and extensible (i.e., custom variants of MCMC samplers can be built). One of the unique characteristics of `NIMBLE` is that it provides several tools for automatically adapting and selecting samplers to handle models with problematic posterior distributions such as the DDM (Nguyen et al., 2019; Turek et al., 2017).

Proposed Research Plan

1. Implement the DDM probability density function in `NIMBLE` based on our existing `C++` implementation.[1]
2. Decide for an existing psychological data set to which the DDM can be applied. For example, we already have the following data sets:
   - Medical decision making with speed and accuracy emphasis manipulation (Trueblood et al., 2018)
   - Perceptual decision making (Ratcliff and Rouder, 1998)

[1] In a MathSys MSc. project from last year, we (Kendal Foster and I) have implemented (in `C++`) and compared existing approximations to the DDM probability density function which contains an infinite sum. Our results showed that a specific variation of the existing methods is fastest. We will use these implementations as the basis for the proposed project.

- Reading and word perception (Wagenmakers et al., 2008)
- Effect of ageing on numerosity discrimination, recognition memory, and lexical decision (Ratcliff et al., 2010)

3. Fit the DDM to the individual-participants data using `NIMBLE`.
4. Fit a hierarchical-variant of the DDM to the data using `NIMBLE`.
5. Compare the `NIMBLE` fit with a fit using `brms` and `Stan`.
6. Extend the model using either contaminant processes (Ratcliff and Tuerlinckx, 2002) or across-trial variabilities of the DDM parameters as part of the graphical model (i.e., the hierarchical structure).

After this, the project could lead to a PhD focussing on evidence accumulation models or other models popular in mathematical psychology (Wixted and Wagenmakers, 2018). For example, one could focus on increasingly complex variants of the DDM without a analytic likelihood function and how to estimate those models in a Bayesian setting (Hawkins et al., 2015; Ratcliff et al., 2016). Alternatively, one could focus on other modern variants of evidence accumulation models and how to extend or estimate them efficiently (Hawkins and Heathcote, 2020). One important goal of a PhD would be to apply an evidence accumulation model to a real-life data set such as from consumer decision making.

Potential Benefits

Efficient hierarchical Bayesian estimation methods are a precursor for using the DDM with real-life data as real-life data is commonly sparse (i.e., few observation on the individual-level). So far applications to such data are rare although large real-life choice and response time data is increasingly being available through internet or browsing behaviour. The ultimate goal is to make such applications more likely and achievable.

References

Perry de Valpine, Daniel Turek, Christopher J. Paciorek, Clifford Anderson-Bergman, Duncan Temple Lang, and Rastislav Bodik. Programming With Models: Writing Statistical Algorithms for General Model Structures With NIMBLE. *Journal of Computational and Graphical Statistics*, 26(2):403–413, April 2017. ISSN 1061-8600. DOI: 10.1080/10618600.2016.1172487.

Christopher Donkin and Scott D Brown. Response Times and Decision-Making. In John T. Wixted and Eric-Jan Wagenmakers, editors, *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, volume 5, pages 349–381. 2018.

David Gunawan, Guy E. Hawkins, Minh-Ngoc Tran, Robert Kohn, and Scott Brown. New Estimation Approaches for the Linear Ballistic Accumulator Model. *arXiv:1806.10089 [stat]*, January 2020.

Guy Hawkins and Andrew Heathcote. Racing Against The Clock: Evidence-Based Vs. Time-Based Decisions. January 2020. DOI: 10.31234/osf.io/m4uh7.

Guy E. Hawkins, Eric-Jan Wagenmakers, Roger Ratcliff, and Scott D. Brown. Discriminating evidence accumulation from urgency signals in speeded decision making. *Journal of Neurophysiology*, 114(1):40–47, July 2015. ISSN 0022-3077, 1522-1598. DOI: 10.1152/jn.00088.2015.

Andrew Heathcote, Yi-Shin Lin, Angus Reynolds, Luke Strickland, Matthew Gretton, and Dora Matzke. Dynamic models of choice. *Behavior Research Methods*, 51(2):961–985, April 2019. ISSN 1554-3528. DOI: 10.3758/s13428-018-1067-y.

Dao Nguyen, Perry de Valpine, Yves Atchade, Daniel Turek, Nicholas Michaud, and Christopher Paciorek. Automatic adaptation of MCMC algorithms. *arXiv:1802.08798 [stat]*, November 2019.

R. Ratcliff. A theory of memory retrieval. *Psychological Review*, 85(2):59–108, 1978. ISSN 0033-295X.

R. Ratcliff and G. McKoon. The diffusion decision model: Theory and data for two-choice decision tasks. *Neural computation*, 20(4):873–922, 2008. ISSN 0899-7667. DOI: 10.1162/neco.2008.12-06-420.

R. Ratcliff and F. Tuerlinckx. Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, 9(3):438–481, 2002.

Roger Ratcliff and Jeffrey N. Rouder. Modeling Response Times for Two-Choice Decisions. *Psychological Science*, 9(5):347–356, 1998. DOI: 10.1111/1467-9280.00067.

Roger Ratcliff, Anjali Thapar, and Gail McKoon. Individual Differences, Aging, and IQ in Two-Choice Tasks. *Cognitive psychology*, 60(3):127–157, May 2010. ISSN 0010-0285. DOI: 10.1016/j.cogpsych.2009.09.001.

Roger Ratcliff, Philip L. Smith, Scott D. Brown, and Gail McKoon. Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, 20 (4):260–281, April 2016. ISSN 1364-6613. DOI: 10.1016/j.tics.2016.01.007.

Jennifer S. Trueblood, William R. Holmes, Adam C. Seegmiller, Jonathan Douds, Margaret Compton, Eszter Szentirmai, Megan Woodruff, Wenrui Huang, Charles Stratton, and Quentin Eichbaum. The impact of speed and bias on

the cognitive processes of experts and novices in medical image decision-making. *Cognitive Research: Principles and Implications*, 3(1):28, July 2018. ISSN 2365-7464. DOI: 10.1186/s41235-018-0119-2.

Daniel Turek, Perry de Valpine, Christopher J. Paciorek, and Clifford Anderson-Bergman. Automated Parameter Blocking for Efficient Markov Chain Monte Carlo Sampling. *Bayesian Analysis*, 12(2):465–490, June 2017. ISSN 1936-0975, 1931-6690. DOI: 10.1214/16-BA1008.

Brandon M. Turner and Trisha Van Zandt. A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology*, 56(2):69–85, April 2012. ISSN 00222496. DOI: 10.1016/j.jmp.2012.02.005.

Brandon M. Turner, Per B. Sederberg, Scott D. Brown, and Mark Steyvers. A method for efficiently sampling from distributions with correlated dimensions. *Psychological Methods*, 18(3):368–384, September 2013. ISSN 1082-989X. DOI: 10.1037/a0032222.

Joachim Vandekerckhove, Francis Tuerlinckx, and Michael D. Lee. Hierarchical diffusion models for two-choice response times. *Psychological Methods*, 16 (1):44–62, 2011. ISSN 1939-1463, 1082-989X. DOI: 10.1037/a0021765.

Eric-Jan Wagenmakers, R. Ratcliff, P. Gomez, and G. McKoon. A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, 58(1):140–159, 2008. ISSN 0749-596X.

John T. Wixted and Eric-Jan Wagenmakers, editors. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience - Vol. 5: Methodology*. John Wiley & Sons, Inc., Hoboken, NJ, USA, March 2018. ISBN 978-1-119-17016-7 978-1-119-17017-4.