Evidence Accumulation Modeling:
Bayesian Estimation using Differential Evolution

Andrew Heathcote (andrew.heathcote@utas.edu.au)
School of Medicine, The University of Tasmania
Social Sciences Building, Sandy Bay, 7005, Tasmania, Australia
Brandon Turner (turner.826@osu.edu)
Department of Psychology, The Ohio State University, Columbus, OH 43227 USA
Scott D. Brown (scott.brown@newcastle.edu.au)
School of Psychology, The University of Newcastle,
Psychology Building, University Avenue, Callaghan, 2308, NSW, Australia

Evidence Accumulation Modeling

Understanding decision making requires a dynamic approach that accounts for the time taken to make choices as well as the choices that are made. The success of the dynamic approach is underpinned by cognitive models, such as the drift-diffusion model (DDM: Ratcliff & McKoon, 2008) and linear ballistic accumulator (LBA: Brown & Heathcote, 2005, 2008), that attribute decisions to an evidence accumulation processes. The ability of these models, and elaborations of them, to account for the speed and accuracy with which people make decisions across a broad range of tasks has led to an increasing number of applications in Cognitive Science and Neuroscience (for recent examples see Cassey, Heathcote & Brown, 2014; Heathcote, Loft & Remington, in press; Mittner et al., 2014; Turner, Van Maanen, Forstmann, in press).
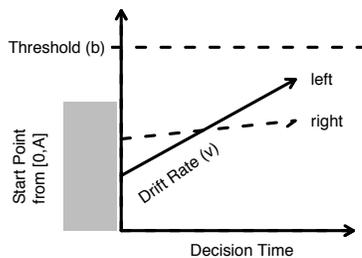

Figure 1. LBA Model for a decision between "left" and "right" responses. Accumulators (arrows) race and the first one to reach threshold determines the choice.

Figure 1 schematically illustrates an LBA model, for a binary ("left" vs. "right") choice. Each choice has its own accumulator that linearly accrues corresponding evidence (illustrated by the arrows in Figure 1), starting from starting points (uniformly distributed over the interval 0-A) that represent random biases from trial to trial. The rate of accumulation (v), which also varies normally from trial to trial, corresponds to the strength of evidence for a choice. The first accumulator to reach its threshold (b) determines the choice. The time for non-decision processes (e.g., stimulus encoding and response selection, t_{er}) is estimated by the difference between the observed response time (RT) and decision time (i.e., the time that evidence in the winning accumulator first equals the threshold). Brown and Heathcote (2008) showed that the LBA is able to account for a comprehensive set of benchmark phenomena in simple choice paradigms, such as speed-accuracy tradeoff (e.g., being more accurate at the cost of longer decision times by raising b). The LBA is easily extended to more than two choices, whereas the DDM, which consists of a single unit with two thresholds and extra sources of noise (in the evidence within trials and uniform trial-to-trial variability in non-decision time), only applies to binary choice.

However, applying models like the LBA and DDM to real data sets can be challenging for a range of reasons: 1) suitable experimental designs are required with sufficient number of observations, experimental control, and manipulations that help to identify model parameters, 2) nonlinear interactions within the models, and only partial observation of the accumulation process (i.e., its end point), cause strong correlations among parameters that make them difficult to estimate, and 3) fitting can be computationally demanding and suffer from problems of numerical instability.

Although excellent estimation packages based on both maximization (Vandekerckhove & Tuerlinckx, 2008; Voss & Voss, 2007) and Bayesian (Vandekerckhove, Tuerlinckx & Lee, 2011; Wiecki, Sofer & Frank, 2013) methods are now available that ease the computational problems for the DDM, they can impose assumptions that do not make them flexible enough for some applications, particularly in the Bayesian setting. This tutorial provides training in a more flexible approach that can be used with the LBA and DDM.

Tutorial Overview

The tutorial is presented by the developers of the LBA model, Scott Brown and Andrew Heathcote, along with Brandon Turner, who with Brown and colleagues proposed using the Differential Evolution algorithm as a way of dealing with the problem of correlated parameters in the Bayesian context (Turner, Sederberg, Brown, & Steyvers, 2013). It teaches attendees to apply Bayesian estimation using DE, with a focus on the LBA model, although the techniques taught can be applied to any evidence accumulation model for which a likelihood can be computed.

Flexibility is obtained by providing attendees with easily modifiable source code in the R language (R Core Team, 2014). This means that some familiarity with the R language is necessary for attendees to best benefit from the tutorial. However, much of the tedious bookkeeping necessary to fit models, as well as graphing to check the results of sampling, is taken care of by convenience functions that users should not have to modify in most standard contexts. Hence, knowledge of R is mainly required to read in data and analyze the outputs of sampling. One the other hand more advanced R users can use these functions as a basis to implement model variants and/or to address non-standard applications.

Likelihood computation uses the newly developed *rtdists* R package (Brown, Gretton, Heathcote & Singmann, 2014), so users do not have to be concerned with the associated mathematical details. This package robustly and efficiently implements likelihood computations for the LBA, including uniform trial-to-trial variability in non-decision time, and the full DDM model with trial-to-trial variability in starting points, mean rate and non-decision time[1].

The tutorial will begin with 30-minute overviews of Bayesian estimation, the LBA model, and DE sampling given, respectively, by Turner, Brown and Heathcote. The remainder of the tutorial will consist of hands on exercises using the suite of R functions provided to simulate data and run sampling for the LBA model, and to examine and interpret outputs. The first session up to lunch will focus on estimation for an individual participant in a design with two conditions. After lunch the idea of hierarchical modeling will be introduced (i.e., estimation of group as well as individual level parameters in a design with a within-subjects factor. If time permits students will be introduced to code provided for more complex designs (e.g., both with and between-subjects factors).

Finally, time will be set aside to discuss with the presenters how to approach the applications of the LBA and DDM model they are interested in making. Advice will be given on the design of appropriate experiments, analyses of existing data and approaches to more advanced designs and models. Students are encouraged to bring their own data sets or proposed designs in order to facilitate discussion.

References

Scott Brown, Matthew Gretton, Andrew Heathcote and Henrik Singmann (2014). rtdists: Response time distributions. R package version 0.2-6. http://CRAN.R-project.org/package=rtdists

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178.

Cassey, P., Heathcote, A., & Brown, S. D. (2014) Brain and Behavior in Decision-Making. PLoS Computational Biology, 10(7), e1003700.

Heathcote, A. (2014). Flexibility and selective influence in evidence accumulation models. Keynote paper presented at the European Mathematical Psychology Group Meeting, July, Tubingen, Germany.

Heathcote, A., Brown, S.D. & Mewhort, D.J.K. (2002). Quantile Maximum Likelihood Estimation of Response Time Distributions. *Psychonomic Bulletin and Review*, *9*, 394-401.

Heathcote, A., Loft, S. & Remington, R. (accepted 31/12/14). Slow down and remember to remember! A delay theory of prospective memory costs, *Psychological Review*

Mittner, M., Boekel, W., Tucker, A., Turner, b., Heathcote, A. & Forstmann, B. U. (2014). When the brain takes a break: A model-based analysis of mind wandering, *Journal of Neuroscience, 34(49),*16286-16295.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Turner, B., Sederberg, P., **Brown,** S.D., & Steyvers, M. (2013) A method for efficiently sampling from distributions with correlated dimensions. *Psychological Methods*, 18(3), 368-384.

Turner, B., Van Maanen, L., & Forstmann, B. (in press). Combining cognitive abstractions with neurophysiology: The Neural Drift Diffusion Model. *Psychological Review*.

Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion Model Analysis with MATLAB: A DMAT Primer. *Behavior Research Methods, 40,* 61-72.

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion Models in Experimental Psychology. *Experimental Psychology; Experimental Psychology*, *60*(6), 385–402.

Voss, A., & Voss, J. (2007). Fast-dm: A Free Program for Efficient Diffusion Model Analysis. *Behavioral Research Methods, 39,* 767-775.Voss, A., & Voss, J. (2007). Fast-dm: A Free Program for Efficient Diffusion Model Analysis. *Behavioral Research Methods, 39,* 767-775.

Wiecki, T.V., Sofer, I. & Frank, M.J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinform*atics, 7:14.

Wabersich, D., & Vandekerckhove, J. (2014). The RWiener package: an R package providing distribution functions for the Wiener diffusion model. *The R Journal, 6,* 49-56.

[1] The likelihood for the DDM with only within-trial variability can be computed efficiently (Wabersich & Vandekerckhove, 2014), but adding trial-to-trial variability requires numerical integration, which can be unstable and slow. Hence, most most previous full-DDM applications have been used quantile-based methods (e.g., Heathcote, Brown & Mewhort, 2002) that are not ideal in the Bayesian context, or have implemented trail-to-trial variability hierarchically (Vandekerckhove et al., 2013), which can result in an overly flexible model (Heathcote, 2014). The *rtdists* package is based on Voss, Nagler and Lerche's (2013) efficient C code, making likelihood-based estimation of the full DDM practical.