Journal of Mathematical Psychology 60 (2014) 58–71

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

A cognitive latent variable model for the simultaneous analysis of behavioral and personality data

Joachim Vandekerckhove *

University of California, Irvine, United States

HIGHLIGHTS

- Cognitive models and latent variable models are combined into a single framework.
- The framework can be applied in a Bayesian inferential context.
- An application uses data from an RT experiment and questionnaires in a single model.

ARTICLE INFO

Article history:
Received 7 September 2013
Received in revised form
13 June 2014
Available online 12 July 2014

Keywords:
Individual differences
Cognitive model
Latent variable
Factor analysis
Data fusion
Diffusion model

ABSTRACT

I describe a *cognitive latent variable model*, a combination of a cognitive model and a latent variable model that can be used to aggregate information regarding cognitive parameters across participants and tasks. The model is ideally suited for uncovering relationships between latent task abilities as they are expressed in experimental paradigms, but can also be used as data fusion tools to connect latent abilities with external covariates from entirely different data sources. An example application deals with the structure of cognitive abilities underlying an executive functioning task and its relation to personality traits.

1. Introduction

1.1. Cognitive psychometrics

Cognitive psychometrics is the term coined by Batchelder (1998) to describe the application of cognitive process models as assessment tools, or, more fundamentally, to apply the psychometrics of individual differences to cognitive process parameters. The practice of combining cognitive measurement models with individual variability, implemented as statistical random effects, serves in the first place to adapt cognitive models to the reality of randomly sampled, noninterchangeable participants (e.g., Batchelder, 2007). As has been pointed out by Estes (1956, 2002), Hamaker (2012), and Heathcote, Brown, and Mewhort (2000), averaging artifacts can lead to biased estimates and errors in inference. More than that, however, the assumption that an individual's process parameters are in fact a random draw from some superordinate population distribution introduces a crucial new aspect to cognitive modeling: The idea that there might be formal structure to be derived from the individual differences researchersoften observe among participants' cognitive model parameters.

Structured individual differences are a critical concept in certain fields of cognitive science. For example, intelligence research is dominated by studies in which individuals are assessed on a variety of tasks, and it is typically observed that participants who score high on one task also score high on other tasks (e.g., Kamphaus, Petoskey, & Morgan, 1997). This covariance is taken to imply that there exists a small set of person-specific abilities that jointly give rise to correlated behavior on the larger set of tasks (a "positive manifold"). An identical approach is often taken in fields such as working memory (e.g., Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000) or executive functioning (e.g., Miyake et al., 2000), where unobserved factors supporting stable differences across individuals are inferred from the correlational pattern between multiple basic tasks. This type of data analysis is widely known as latent variable modeling (Bartholomew, Knott, & Moustaki, 2011; Skrondal & Rabe-Hesketh, 2004).

* Correspondence to: University of California, 2324 SBSG, Irvine, CA 92617-5100, United States.
E-mail address: joachim@uci.edu.

http://dx.doi.org/10.1016/j.jmp.2014.06.004
0022-2496/

Importantly, the interpretability and usefulness of the results of such analyses depend on the interpretability of the quantities measured in the basic tasks. If each score in a given set of tasks can reasonably be thought to tap intelligence, then it is valid to conclude that the inferred latent factors relate to intelligence as well. If, on the other hand, scores in the basic tasks are nonlinear amalgams of more elementary variables, interpretation of the latent factors is complicated. Cognitive models serve to decompose such complex data into interpretable parameters. The modeling strategy proposed in this paper involves – within a single model – a latent variable structure built on top of a cognitive process model, to allow inference of latent variables that have cognitive interpretations.

1.2. A qualitatively different type of conclusion

When latent variable models are combined with cognitive models to form a *cognitive latent variable model* (CLVM), this affords a qualitatively different type of conclusion from either classical psychometrics or classical cognitive modeling. For example, using a cognitive model with a parameter interpreted as *speed of information processing* (e.g., the drift rate in a diffusion model Ratcliff, 1978), a CLVM permits inferences about unobserved variables that contribute to the total rate of information processing in a particular task. A conventional psychometric model would not permit such process-based conclusions, whereas a conventional cognitive model would not be equipped to infer higher-order latent properties.

Combining cognitive models with latent variable models allows us to bridge the gap between experimental and individual-differences research—a long-standing issue in psychology since Cronbach's (1957) lament that the science is split across two disparate disciplines, reiterated more recently by Borsboom (2006). It is the aim of the present paper to present an example of a CLVM, a formal model that extends the logic of cognitive psychometrics to include latent variable structures.

The structure of the paper is as follows. The next section will introduce two components of the CLVM: the diffusion model as a cognitive model of choice response time data and the factor analysis model as a measurement model to tie multiple tasks together. This section will also introduce some required notation. The section after that will focus on properties of the integrative CLVM. After that, a short section will be devoted to the relevant details of Bayesian inference and model selection. Finally, a section will provide detail regarding the application of the CLVM in the field of emotion psychology.

2. Diffusion models for two-choice RT

The data level of this CLVM consists of a probabilistic representation of data as they are predicted by a particular cognitive model—the sampling scheme of the data. The cognitive model used here is the diffusion model for two-choice RT (Stone, 1960), which has been very popular in cognitive science (Wagenmakers, 2009), with applications ranging from memory (Ratcliff, 1978) and low-level perception (Ratcliff & Rouder, 1998) to semantic cognition (Vandekerckhove, Verheyen, & Tuerlinckx, 2010) and emotion psychology (Pe, Vandekerckhove, & Kuppens, 2013; White, Ratcliff, Vasey, & McKoon, 2009). The diffusion model is based on the principle of *sequential accumulation of information*—it assumes that a decision making system samples small units of information, sequentially over time, from whatever stimulus to which it was exposed. These sampled units of evidence are aggregated with information already accumulated. After each accretion step, the system evaluates whether the total amount of evidence warrants the making of a decision. If so, the process ends and a response is executed. This accumulation process is the fundamental assumption – the "central dogma" – of a broad and highly successful class of sequential sampling models for RT.

More specifically, the process assumptions of the diffusion model are that a single evidence counter accumulates towards one of two decision boundaries, with a starting point that may be closer to one boundary than the other. Fig. 1 illustrates the process. Given the freedom of two decision bounds, the model can account for two distinct types of bias in the response process. In addition to biased processing of information (which is reflected in the average rate of evidence accumulation, a parameter called the *drift rate*, δ), the diffusion model allows for an a-priori bias that is prior to and independent of the information accumulation process (here parameterized as a proportion, so that a *bias* $\beta = 0.5$ implies a-priori indifference). The distance between the decision bounds (known as the *boundary separation* α) performs a separate, interesting task in the diffusion process. Bounds that are close together lead to fast decisions that are largely independent from the information contained in the stimulus (i.e., close to chance level), whereas distant bounds lead to slow response processes whose outcome is mostly determined by the direction of the accumulation process (i.e., if δ is positive and α is high, the upper boundary is likely to be hit). This parameter hence captures the well-known *speed–accuracy trade-off*. The fourth and final parameter of the diffusion model is the *nondecision time* τ. This shift parameter determines the leading edge of the latency distribution, and is typically interpreted as the sum duration of all non-decision processes (and it is additionally assumed that these processes are independent of and serial to the decision process).

The PDF of the Wiener diffusion model is bivariate (with one dimension for the latency and one for the binary choice); its analytical form also contains an infinite sum and the latency distribution can therefore at best be approximated:

$$\begin{cases} p(t, x=0|\alpha,\beta,\tau,\delta) = \dfrac{\pi}{\alpha^2} e^{-\frac{1}{2}\left(2\alpha\beta\delta+\delta^2(t-\tau)\right)} \\ \qquad\qquad \times \displaystyle\sum_{k=1}^{+\infty}\left[k\sin(\pi k\beta) e^{-\frac{1}{2}\frac{k^2\pi^2}{\alpha^2}(t-\tau)}\right] \\ p(t, x=1|\alpha,\beta,\tau,\delta) = p(t, x=0|\alpha, 1-\beta, \tau, -\delta). \end{cases} \quad (1)$$

Fortunately, efficient methods for the computation of the Wiener diffusion model density and distribution functions exist (Blurton, Kesselmeier, & Gondan, 2012; Navarro & Fuss, 2009, for the CDF and PDF, respectively), making it a highly tractable model. Eq. (1) lacks a *diffusion coefficient* parameter, which is sometimes used to scale the evidence dimension (and typically denoted s); the coefficient does not appear because it will be set to 1 in all applications, and it cancels out everywhere.

Fig. 2 shows a graphical model representation of a Wiener diffusion model for a data set where P participants do a task with T conditions and I trials in each condition. For conciseness, y denotes a choice RT pair (t, x). The equations to the right of the diagram list the distributional assumptions of the model, including some example priors.

It is important to note that this data model can serve a dual purpose for researchers in psychology. On the one hand, researchers can decide to buy in to the assumptions of the model—taking the process as given and drawing conclusions that may hinge on the accuracy of these assumptions. For this particular cognitive model, the literature contains reports of experimental manipulations that selectively affect model parameters, lending some credibility to the process assumptions (e.g., Voss, Rothermund, & Voss, 2004). However, the model would remain useful even if one is unwilling to buy in to the exact process—by taking the model as a convenient data level that captures the shape of the data and serves strictly as a parsimonious description.