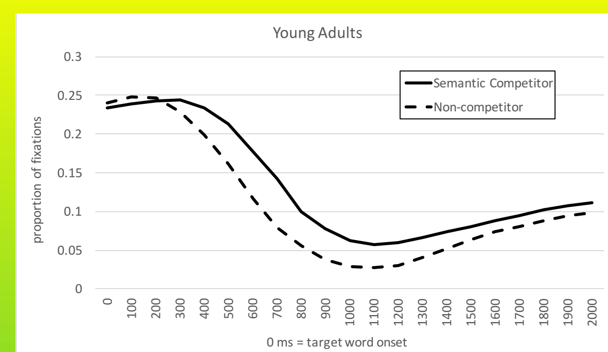
The dynamic generalized linear mixed effect model: Modeling intensive binary time-series data from the visual-world eye-tracking paradigm with GLMM with crossed random effects

Sarah Brown-Schmidt and Sun-Joo Cho (Vanderbilt University)
sarahbrownschmidt@gmail.com

Background

The Visual World eye-tracking Paradigm (Tanenhaus et al., 1995) produces intensive categorical time-series data indicating which of one or more object categories is being fixated at each time point.



Why the need for the dGLMM?

The literature lacked a model specification that could:

- Analyze the full time course of data without aggregation over time-points, trials, items, and/ or participants
- Analyze the data in categorical (e.g., binary) form
- Model crossed random effects
- Model temporal (trend) effects
- Model autocorrelation (AR) among adjacent time-points (estimated to be very high in VWP data with *r* values typically exceeding .9 for 10ms time-bins)

Implementation

How to format your data to implement a dGLMM:

- Pick analysis time-window a-priori
- Pick temporal grain size (e.g., 10ms)
- Code fixation data in binary form (e.g., target vs. non-target)
- Save the data in long form
- Calculate the order of the AR: AR(1), or AR(2), etc.
- Add the AR to the dataset (will require baseline time-points to calculate AR for first time-point)
- Remove starting time-point

trial ID	subject ID	item	fixed effect 1	fixed effect 2	time	target fixations	AR(1)
273	1	balloon	0.5	0.5	180	0	NA
273	1	balloon	0.5	0.5	190	0	0
273	1	balloon	0.5	0.5	200	0	0
273	1	balloon	0.5	0.5	210	0	0
273	1	balloon	0.5	0.5	220	0	0
273	1	balloon	0.5	0.5	230	1	0
273	1	balloon	0.5	0.5	240	1	1
273	1	balloon	0.5	0.5	250	1	1
273	1	balloon	0.5	0.5	260	1	1
273	1	balloon	0.5	0.5	270	1	1
273	1	balloon	0.5	0.5	280	1	1
273	1	balloon	0.5	0.5	290	1	1
273	1	balloon	0.5	0.5	300	1	1

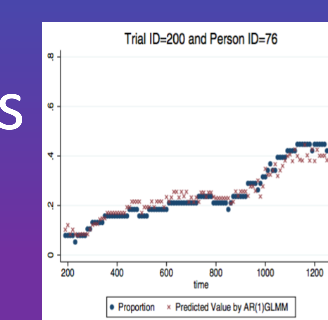
- **Fit the model in glmer:**

mymodel ← glmer (y ~ 1 + AR1+ condition + (1 | trialID) + (1+AR1 | subjectID) + (1+AR1 | item), family = binomial, data = mydata)

Benefits

Why might you want to use dGLMM?

- Test hypotheses about overall level of activation of, e.g., the target across conditions, while modeling:
  - crossed random effects
  - temporal (trend) effects
  - unaggregated data
- Ignoring AR can result in underestimated SEs and increased type-1 error rate
  - *bonus*: can test if AR varies systematically across, e.g. persons or conditions
- The dGLMM provides very good fits to VWP data



Conclusions

- The dGLMM (Cho et al. 2018) offers a way to model intensive categorical time-series data with crossed random effects, temporal (trend) effects, and captures autocorrelation present in the data.
- Extensions of the dGLMM can additionally model spatial information (Cho et al., forthcoming), which can vary across persons and items.
- Visit poster for application of dGLMM: #D9 On the necessity of hippocampus in lexical-semantic mapping in language processing.

References:

Cho, S.-J., Brown-Schmidt, S., & Lee, W.-y. (2018). Autoregressive generalized linear mixed effect models with crossed random effects: An application to intensive binary time series eye-tracking data. *Psychometrika, 83*, 751-771. **Cho, S.-J., Brown-Schmidt, S., Naveiras, M., & De Boeck, P. (forthcoming).** A dynamic generalized mixed effect model for intensive binary temporal-spatio data from an eye tracking technique. Maryland Assessment Research Center (MARC). **Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995).** Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632-1634.

Acknowledgement: NSF BCS 15-56700 to SBS.

Look for example data and code at sarahbrownschmidt.com/cv/