

## Statistics at ZAS (1):

### Hypothesis testing using linear mixed effects models

This workshop is devoted to a detailed discussion of several issues relating to linear mixed effects models (LMMs), which are widely used in linguistics, psychology, and cognitive science. An important recent paper in this context is Barr et al. (2013).

Barr et al. (2013) have raised a lot of awareness about random effects in LMMs. Their main argument is that LMMs “generalize best when they include the maximal random effects structure justified by the design” (Barr et al. 2013:255). The paper seeks to provide a simple procedure for defining LMMs for data sets typically seen in psycholinguistic work. In this workshop, we will attempt to unpack the main ideas of the Barr et al. paper, and then move on to discussing the recent paper by Bates et al. (2015), which addresses the question of how to proceed when there is not enough data to recover accurate estimates of the true parameters in the random effects part of the model. Finally, in the context of LMMs we will discuss the more general question of what the purpose of a statistical model actually is. Psychologists and statisticians have, it seems, very different answers to this question. A major goal of this workshop is thus to understand what statisticians have to say about statistical models used for hypothesis testing.

The workshop will consist of two parts. We will start with a theoretically oriented lecture followed by a discussion which will focus on the variance-covariance matrices of the random effects in maximal versus non-maximal LMMs. In the second part of the workshop, we will take up some practical examples.

The workshop will be led by

**Prof. Shravan Vasishth** from the University of Potsdam

<http://www.ling.uni-potsdam.de/~vasishth/>

As preparation, it is recommended to:

- 1) view the ESSLLI 2015 lectures (week 1 only) by Shravan Vasishth and do the automatically graded exercises (<http://www.ling.uni-potsdam.de/~vasishth/statistics/ESSLLI2015Vasishth.html>)
- 2) read the articles by Barr et al. (2013) [<http://idiom.ucsd.edu/~rlevy/papers/barr-etal-2013-jml.pdf>] and Bates et al. (2015) [<http://arxiv.org/pdf/1506.04967v1.pdf>]
- 3) install the latest version of R, lme4 (<https://github.com/lme4/lme4/>), and RePsychLing (<https://github.com/dmbates/RePsychLing>).
- 4) read (optionally):
  - (a) The lecture notes from the Potsdam summer semester linear modeling class (MSc Cognitive Systems) which provides technical background on linear models: <https://github.com/vasishth/LM>,
  - (b) The tutorial on fitting Bayesian LMMs using Stan, which shows how over-parameterized models can be defined in a Bayesian setting: <http://www.ling.uni-potsdam.de/~vasishth/statistics/BayesLMMs.html>

The workshop will take place on October 7<sup>th</sup> from 9:30 until 17:00 at the Centre for General Linguistics (ZAS) in Berlin (Schützenstr. 18, D-10117) in seminar room **303**.

Since the number of places is limited, we kindly ask you to register yourself by sending an email to [statistics-zas@zas.gwz-berlin.de](mailto:statistics-zas@zas.gwz-berlin.de) to reserve a spot by October 5<sup>th</sup>.

Marzena Zygis  
Sergio Quiroz

References:

Barr, D. J., Levy, R., Scheepers, Ch. & H. J. Tily (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3):255-278.

Bates, D., Kliegl, R., Vasishth, S. and Baayen, H. (2015). Parsimonious mixed models. ArXiv e-print; under revision.