Basic Bayes for Linguists
Aleksandra Ćwiek

Abstract

This course will show you that Bayesian data analysis doesn't have to be intimidating. In fact, once you get the hang of it, you'll find that Bayesian inference is often easier to understand and evaluate than frequentist approaches – and it's more robust, too.

We'll focus on the kinds of questions and data structures we encounter in linguistics research. You'll develop an intuitive grasp of how Bayesian inference works: how prior knowledge combines with your data to give you posterior beliefs, and how to turn that understanding into a practical workflow for estimation, quantifying uncertainty, and making predictions.

Topics include likelihoods and priors, interpreting posteriors and credible intervals, model checking through posterior predictive checks, and model comparison using leave-one-out validation. We'll then work with regression models and multilevel (hierarchical) structures typical in linguistics: group-level effects by participants, items, or languages, and handling missing data. Along the way, we'll discuss when model comparison makes sense, how to check whether your results are sensitive to your priors, and how Bayesian thinking differs from the "p-values-first" mindset.

You should have some experience with R (ideally also with statistical modeling). We'll use brms as an interface to Stan, learn to diagnose convergence and sampling issues, and practice presenting results clearly. By the end, you'll be able to specify, fit, check, and report Bayesian models for common outcome types – continuous variables, binary data, counts, and ratings – and make informed modeling decisions for your own research