Tutorial: Doing Bayesian Data Analysis with R and BUGS

John K. Kruschke (kruschke@indiana.edu)
Department of Psychological and Brain Sciences, Indiana University
1101 E. 10th St., Bloomington, IN USA

Before arriving, install free software and get more information from this web site:
http://www.indiana.edu/∼jkkteach/CogSci2011Tutorial.html

Keywords: Data analysis (Bayesian); Statistics (Bayesian); Markov chain Monte Carlo; Bayesian models; Hierarchical models

An *introduction* to *doing* Bayesian data analysis

This full-day tutorial shows you how to do Bayesian data analysis, hands on. The software is free. The intended audience is graduate students and other researchers who want a ground-floor introduction to Bayesian data analysis. No mathematical expertise is presumed. The full-day tutorial progresses through the following topics.

9:00-10:30 Bayes' Rule, Grid Approximation, and R. We start with the basics of conditional probabilities, the meaning of Bayes' rule, and simple examples of Bayes' rule graphically illustrated with grid approximation in the programming language R.

11:00-12:30 MCMC and BUGS; Linear Regression. We explain the idea of approximating distributions by large representative samples, and Markov chain Monte Carlo (MCMC) methods for generating them. The BUGS language is introduced and used to do Bayesian linear regression.

1:30-3:00 Hierarchical Models and Model Comparison. Bayesian methods and the BUGS language make hierarchical modeling straight forward. Hierarchical models are tremendously useful for analyzing individual differences, repeated measures, and structural constraints across conditions. Model comparison is a case of hierarchical modeling.

3:30-5:00 Bayesian ANOVA; Power Analysis. We use *hierarchical* analysis of variance (ANOVA) with Bayesian parameter estimates, for rich and flexible inferences about differences between groups. We conclude with a brief look at power analysis from a Bayesian perspective.

Concepts and methods of Bayesian data analysis *transfer* to Bayesian models of cognition

Bayesian data analysis uses generic descriptive models such as linear regression, without any assertions about the pro-

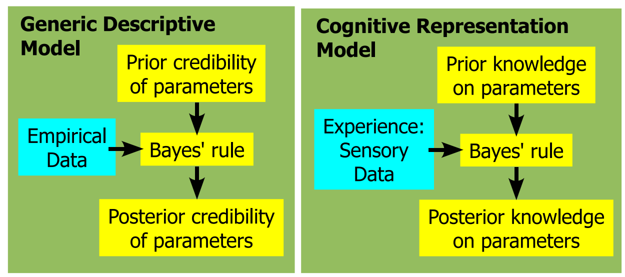

Figure 1: Concepts and methods of Bayesian data analysis (left) transfer to Bayesian models of mind (right), but Bayesian data analysis with generic descriptive models will be useful even when specific Bayesian models of mind fail to mimic behavior.

cesses that generated the data. Bayesian methods infer credible values of parameters in the descriptive models, such as credible slopes and intercepts in linear regression, as suggested in the left side of Figure 1.

Because the Bayesian approach to inference is the normative approach, some cognitive scientists posit that cognitive processing itself is based on Bayesian inference by the mind, as suggested in the right side of Figure 1.When you learn about concepts and methods of Bayesian data analysis, it is easier to understand Bayesian models of mind, because the concepts transfer directly. But Bayesian data analysis will always be useful, even if particular Bayesian models of mind fail to accurately mimic cognition.

Why go Bayesian?

Modern Bayesian methods are the best approach to empirical data analysis because Bayesian methods yield richer inferences than traditional methods and without use of ill-defined p values. Sciences from astronomy to zoology are changing from 20th-century null-hypothesis significance testing to Bayesian data analysis. Figure 2 (humorously) suggests this trend.

Bayesian data analysis delivers many practical benefits:

- Bayesian methods permit model flexibility and appropriateness: Hierarchical models can be built easily to suit the design of the experiment and the type of data measured. Such models can be easily extended to capture individual

Before arriving, install free software and get more information from this web site:
`http://www.indiana.edu/∼jkkteach/CogSci2011Tutorial.html`

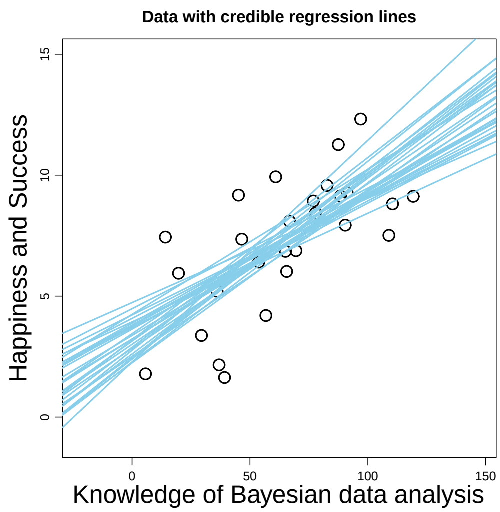

Figure 2: Bayesian linear regression reveals many credible lines, instead of a single "best" line. Also revealed are the inherent trade-offs in estimated slope and intercept: When the slope is steeper, the intercept is lower.

differences, group differences, repeated measures, and for various types of data.

- Bayesian methods reveal credibilities of all *combinations* of parameter values, unlike traditional analysis which has only point estimates. As a simple example, Figure 2 demonstrates that Bayesian linear regression reveals correlations in estimated values of slope and intercept. Knowledge of these trade-offs can be useful in applications involving multiple correlated predictors and other realistic situations.

- Bayesian methods encourage thorough data analysis including multiple comparisons, because there are no penalizing "corrections" for multiple comparisons as in p-value based decisions. Bayesian methods instead create rational shrinkage informed by the data.

- Model flexibility allows conceptual transition from generic descriptive models to domain-specific models wherein parameters serve psychometric purposes.

- Bayesian methods permit different sample sizes in different groups, and different sample sizes per subject, unlike traditional ANOVA which has troubles with unbalanced designs.

- Bayesian methods allow data collection to stop at any time, unlike p-value based decisions that require a pre-set stopping criterion, such as fixed sample size, and no "peeking" at the data.

- Bayesian hypothesis testing permits a principled way to assess evidence in favor of a null hypothesis, unlike NHST.

- Bayesian methods allow cumulative science and use of prior knowledge for leveraged inference when data are sparse, unlike traditional methods.

- Power and replication probability are straight forward to estimate with Bayesian methods, but difficult to assess in p-value based methods.

- 20th century methods, based on p values, have numerous deep problems that are avoided with Bayesian methods.

The tutorial will not have time to address all the points listed above, but will illustrate many of them with examples from linear regression and ANOVA. For a brief discussion of several benefits of Bayesian data analysis, along with a worked example, and an emphasis that Bayesian data analysis is not Bayesian modeling of mind, see Kruschke (2010b). For a lengthier exposition that explains one of the primary pitfalls of null hypothesis significance testing and has a discussion of Bayesian null hypothesis testing, along with different examples, see Kruschke (2010a). A thorough introduction to these issues is presented in Kruschke (2011), but the tutorial presents them in a streamlined, interactive, hands-on format.

The instructor

John Kruschke is five-time winner of Teaching Excellence Recognition Awards from Indiana University, where he is Professor of Psychological and Brain Sciences, and Adjunct Professor of Statistics. In addition to teaching courses for over 20 years, he has presented numerous well-received workshops on Bayesian data analysis. His research interests include the science of morality, applications of Bayesian methods to adaptive teaching and learning, and models of attention in learning, which he has developed in both connectionist and Bayesian formalisms. He received a Troland Research Award from the National Academy of Sciences. He chaired the Cognitive Science Conference in 1992.

References

Kruschke, J. K. (2010a). Bayesian data analysis. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(5), 658–676.

Kruschke, J. K. (2010b). What to believe: Bayesian methods for data analysis. *Trends in Cognitive Sciences*, *14*(7), 293–300.

Kruschke, J. K. (2011). *Doing Bayesian data analysis: A tutorial with R and BUGS*. Burlington, MA: Academic Press / Elsevier.