



Editorial: Bayesian methods for advancing psychological science

Joachim Vandekerckhove¹ · Jeffrey N. Rouder^{1,2} · John K. Kruschke³

Published online: 15 February 2018
© Psychonomic Society, Inc. 2018

Keywords Bayesian statistics · Bayesian inference and parameter estimation · Evidence · New statistics

The simple act of deciding which among competing theories is most likely—or which is most supported by the data—is the most basic goal of empirical science, but the fact that it has a canonical solution in probability theory is seemingly poorly appreciated. It is likely that this lack of appreciation is not for want of interest in the scientific community; rather, it is suspected that many scientists hold misconceptions about statistical methods. Indeed, many psychologists attach false probabilistic interpretations to the outcomes of classical statistical procedures (p values, rejected null hypotheses, confidence intervals, and the like; Gigerenzer, 1998; Hoekstra, Morey, Rouder, & Wagenmakers, 2014; Oakes, 1986). Because the false belief that classical methods provide the probabilistic quantities that scientists need is so widespread, researchers may be poorly motivated to abandon these practices.

The American Statistical Association recently published an unusual warning against inference based on p values (Wasserstein & Lazar, 2016). Unfortunately, their cautionary message did not conclude with a consensus recommendation regarding best-practice alternatives, leaving something of a recommendation gap for applied researchers.

In psychological science, however, a replacement had already been suggested in the form of the “New Statistics” (Cumming, 2014)—a set of methods that focus on effect

size estimation, precision, and meta-analysis, and that would forgo the practice of ritualistic null hypothesis testing and the use of the maligned p value. However, because the New Statistics’ recommendations regarding inference are based on the same flawed logic as the thoughtless application of p values, they are subject to the same misconceptions and are lacking in the same department. It is not clear how to interpret effect size estimates without also knowing the uncertainty of the estimate (and despite common misconceptions, confidence intervals do not measure uncertainty; Morey, Hoekstra, Rouder, Lee, & Wagenmakers, 2016), nor is it clear how to decide which among competing theories is most supported by data.

In this special issue of *Psychonomic Bulletin & Review*, we review a different set of methods and principles, now based on the theory of probability and its deterministic sibling, formal logic (Jaynes, 2003; Jeffreys, 1939). The aim of the special issue is to provide and recommend this collection of statistical tools that derives from probability theory: Bayesian statistics.

Overview of the special issue on Bayesian inference

The special issue is divided into four sections. The first section is a coordinated five-part introduction that starts from the most basic concepts and works up to the general structure of complex problems and to contemporary issues. The second section is a selection of advanced topics covered in-depth by some of the world’s leading experts on statistical inference in psychology. The third section is an extensive collection of teaching resources, reading lists, and strong arguments for the use of Bayesian methods at the expense of classical methods. The final section contains a number of applications of advanced Bayesian analyses that provide an idea of the wide reach of Bayesian methods for psychological science.

The Editors are grateful for the efforts of all contributing authors and reviewers.

JV was supported by National Science Foundation grants #1230118 and #1534472.

✉ Joachim Vandekerckhove
joachim@uci.edu

¹ University of California, Irvine, CA, USA

² University of Missouri, Columbia, MO, USA

³ Indiana University, Indianapolis, IN, USA

Section I: Bayesian inference for psychology The special issue opens with *Introduction to Bayesian inference for psychology*, in which Etz & Vandekerckhove describe the foundations of Bayesian inference. It is illustrated how all aspects of Bayesian statistics can be brought back to the most basic rules of probability, and that Bayesian statistics is nothing more nor less than the systematic application of probability theory to problems that involve uncertainty. With seven worked examples (the seventh split into two parts) of gradually increasing scope and sophistication, the first paper covers a variety of possible practical scenarios, ranging from simple diagnosis to parameter estimation, model selection, and the calculation of strength of evidence.

Wagenmakers et al. continue in *Part I: Theoretical advantages and practical ramifications* by illustration the added value of Bayesian methods, with a focus on its desirable theoretical and practical aspects. Then, in *Part II: Example applications with JASP*, Wagenmakers et al. showcase JASP: free software that can be used to perform the statistical analyses that are most common in psychology, and that can execute them in both a classical and Bayesian way. One of the goals of the JASP project is to provide psychologists with free software that can fulfill all the basic statistical needs that are now met by expensive commercial software. At the time of writing, JASP allows standard analyses like *t* tests, analysis of variance, regression analysis, and analysis of contingency tables.

However, the full power of Bayesian statistics comes to light in its ability to work seamlessly with far more complex statistical models. As a science matures from the mere description of empirical effects to predication and explanation of patterns, more detailed formal models gain center stage. In *Part III: Parameter estimation in nonstandard models*, Matzke et al. discuss the nature of formal models, the requirements of their construction, and how to implement models of high complexity in the modern statistical software packages WinBUGS (Lunn et al., 2000), JAGS (Plummer, 2003), and Stan (Stan Development Team, 2013).

Rounding out the section, Rouder et al. in *Part IV: Parameter estimation and Bayes factors* discuss the fraught issue of estimation-versus-testing. The paper illustrates that the two tasks are one and the same in Bayesian statistics, and that the distinction in practice is not a distinction of method but of how hypotheses are translated from verbal to formal statements. It is an important feature of Bayesian methods that it matters exactly which question one is asking of the data (and sometimes it requires careful thought to assess precisely what question is asked by a particular model), but that once the question is clearly posed, the solution is unambiguous and inescapable.

Additional tutorial articles are provided in Section III, *Learning and Teaching*.

Section II: advanced topics The *Advanced Topics* section covers three important issues that go beyond the off-the-shelf use of statistical analysis. In *Determining informative priors for cognitive models*, Lee & Vanpaemel highlight the sizable advantages that prior information can bring to the data analyst become cognitive modeler. After establishing that priors, just like every other part of a model, are merely codified assumptions that aid in constraining the possible conclusions from data, these authors go on to illustrate some of the different sources of information that cognitive modelers can use to specify priors.

In classical inference, the computation of *p* values and confidence intervals can be corrupted by researchers “peeking” at their data and making a data-dependent decision to stop or continue collecting data (a.k.a. “optional stopping”). Counterintuitively, these classical quantities are invalidated if the researcher makes this forbidden choice, which is part of why it is a recommended practice to pre-register one’s data collection intentions so reviewers can confirm that a well-defined data collection plan was followed. In contrast, Bayesian analyses are not in general invalidated by “peeking” at data and so the use for sample size planning and power analysis is somewhat diminished. Nevertheless, for logistical reasons (planning time and money investments), it is sometimes useful to calculate ahead of time how many participants a study is likely to need before it yields some minimal or criterial amount of evidence. In *Bayes factor design analysis: Planning for compelling evidence*, Schoenbrodt & Wagenmakers provide exactly that.

Finally, there arise occasions where even the most sophisticated general-purpose software will not meet the needs of the expert cognitive modeler. On such occasions, one may want to implement a custom Bayesian computation engine. In *A simple introduction to Markov chain Monte-Carlo sampling*, van Ravenzwaaij et al. describe the basics of sampling-based algorithms, and with examples and provided code, illustrate how to construct a custom algorithm for Bayesian computation.

Section III: learning and teaching Four articles make up the *Learning and Teaching* section. The goal of this section is to collect the most accessible, self-paced learning resources for an engaged novice.

One telling feature of Bayesian methods is how it tends to accord with human intuitions about evidence and knowledge. While it is of course the mathematical underpinnings that support the use of these methods, the intuitive nature of Bayesian inference is a great advantage for novice learners. In *Bayesian data analysis for newcomers*, Kruschke & Liddell cover the basic foundations of Bayesian methods using examples that emphasize this intuitive nature of probabilistic inference. The paper

includes discussion of such topics as the use of priors and limitations of statistical decision-making in general.

With *The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and planning from a Bayesian perspective*, Kruschke & Liddell lay out a broad and comprehensive case for Bayesian statistics as a better fit for the goals of the aforementioned New Statistics (Cumming, 2014). Contrasting the Bayesian approach with the New Statistics—which these authors discuss at length—Kruschke & Liddell review approaches to estimation and testing and discuss Bayesian power analysis and meta-analysis.

How to become a Bayesian in eight easy steps is notable in part because it is an entirely student-contributed paper. Etz et al. thoroughly review a selection of eight basic works (four theoretical, four practical) that together cover the bases of Bayesian methods. The article also includes a *Further Reading* appendix that briefly describes an additional 32 sources that are arranged by difficulty and theoretical versus applied focus.

The fourth and final paper in the section for teaching resources is *Four reasons to prefer Bayesian analyses over significance testing* by Dienes & McLatchie. While our intention for this special issue was to introduce Bayesian statistics without spending too many words on the comparison with classical (“orthodox”) statistics, we did feel that a paper focusing on this issue was called for. One reason why this is important is that it is likely that widespread misconceptions about classical methods have made it seem to researchers that their staple methods have the desirable properties of Bayesian statistics that are, in fact, missing. Dienes & McLatchie present a selection of realistic scenarios that illustrate how classical and Bayesian methods may agree or disagree, proving that the attractive properties of Bayesian inference are often missing in classical analyses.

Section IV: Bayesian methods in action The concluding section contains a selection of fully worked examples of Bayesian analyses. Three powerful examples were chosen to showcase the broad applicability of the unifying Bayesian framework.

The first paper, *Fitting growth curve models in the Bayesian framework* by Oravecz & Muth, provides an example of a longitudinal analysis using growth models. While not a new type of model, it is a framework that is likely to gain prominence as more psychologists focus on the interplay of cognitive, behavioral, affective, and physiological processes that unfold in real time and whose joint dynamics are of theoretical interest.

In a similar vein, methods for dimension reduction have become increasingly useful in the era of Big Data. While psychological data that is massive in volume is

still relatively rare, many researchers now collect multiple modes of data simultaneously and are interested in uncovering low-dimensional structures that explain their observed covariance. In *Bayesian latent variable models for the analysis of experimental psychology data*, Merkle & Wang give an example of an experimental data set whose various measures are jointly analyzed in a Bayesian latent variable model.

The final section of the special issue is rounded out by *Sensitivity to the prototype in children with high-functioning autism spectrum disorder: An example of Bayesian cognitive psychometrics* by Voorspoels et al.. Cognitive psychometrics is the application of cognitive models as measurement tools—here in a clinical context. The practice of cognitive psychometrics involves the construction of often complex nonlinear random-effects models, which are typically intractable in a classical context but pose no unique challenges in the Bayesian framework.

Digital coverage

As part of our efforts to make our introductions to Bayesian methods as widely accessible as possible, we have worked with the Psychonomic Society’s Digital Content Editor, who generously offered to host a digital event relating to this special issue. On <http://bit.ly/BayesInPsych>, interested readers may find further expert commentary and web links to more content.

Additionally, the editors have set up a social media help desk (<http://bit.ly/BayesGroup>) where questions regarding Bayesian methods and Bayesian inference, especially as they are relevant for psychological scientists, are welcomed. These two digital resources are likely to expand in the future to cover new developments in the dissemination and implementation of Bayesian inference for psychology.

Finally, we have worked to make many of the contributions to the special issue freely available online. The full text of many articles is freely available via <https://osf.io/2es64>. Here, too, development of these materials is ongoing, for example with the gradual addition of exercises and learning goals for self-teaching or classroom use.

References

- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25, 7–29.
- Dienes, Z., & McLatchie, N. (this issue). Four reasons to prefer Bayesian analyses over significance testing. *Psychonomic Bulletin and Review*.
- Etz, A., Gronau, Q. F., Dablander, F., Edelsbrunner, P. A., & Baribault, B. (this issue). How to become a Bayesian in eight easy steps: An annotated reading list. *Psychonomic Bulletin and Review*.

- Etz, A., & Vandekerckhove, J. (this issue). Introduction to Bayesian inference for psychology. *Psychonomic Bulletin and Review*.
- Gigerenzer, G. (1998). We need statistical thinking, not statistical rituals. *Behavioral and Brain Sciences*, *21*, 199–200.
- Hoekstra, R., Morey, R. D., Rouder, J. N., & Wagenmakers, E. J. (2014). Robust misinterpretation of confidence intervals. *Psychonomic Bulletin & Review*, *21*(5), 1157–1164.
- Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge: Cambridge University Press.
- Jeffreys, H. (1939). *Theory of probability, 1st edn*. Oxford: Oxford University Press.
- Kruschke, J. K., & Liddell, T. M. (this issue). Bayesian data analysis for newcomers. *Psychonomic Bulletin and Review*.
- Kruschke, J. K., & Liddell, T. M. (this issue). The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and planning from a Bayesian perspective. *Psychonomic Bulletin and Review*.
- Lee, M. D., & Vanpaemel, W. (this issue). Determining informative priors for cognitive models. *Psychonomic Bulletin and Review*.
- Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS – a Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing*, *10*, 325–337.
- Matzke, D., Boehm, U., & Vandekerckhove, J. (this issue). Bayesian inference for psychology, part III: Parameter estimation in nonstandard models. *Psychonomic Bulletin and Review*.
- Merkle, E., & Wang, T. (this issue). Bayesian latent variable models for the analysis of experimental psychology data. *Psychonomic Bulletin and Review*.
- Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E. J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*, *23*, 103–123.
- Oakes, M. (1986). *Statistical inference: a commentary for the social and behavioral sciences*. New York: Wiley.
- Oravecz, Z., & Muth, C. (this issue). Fitting growth curve models in the Bayesian framework. *Psychonomic Bulletin and Review*.
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In Hornik, K., Leisch, F., & Zeileis, A. (Eds.) *Proceedings of the 3rd international workshop on distributed statistical computing*, Vienna, Austria.
- Rouder, J. N., Haaf, J., & Vandekerckhove, J. (this issue). Bayesian inference for psychology, part IV: Parameter estimation and Bayes factors. *Psychonomic Bulletin and Review*.
- Schoenbrodt, F., & Wagenmakers, E. J. (this issue). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin and Review*.
- Stan Development Team (2013). Stan: a c++ library for probability and sampling, version 1.1. Retrieved from: <http://mc-stan.org/>.
- van Ravenzwaaij, D., Cassey, P., & Brown, S. (this issue). A simple introduction to Markov chain Monte-Carlo sampling. *Psychonomic Bulletin and Review*.
- Voorspoels, W., Rutten, F., Bartlema, A., Tuerlinckx, F., & Vanpaemel, W. (this issue). Sensitivity to the prototype in children with high-functioning autism spectrum disorder: an example of Bayesian cognitive psychometrics. *Psychonomic Bulletin and Review*.
- Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ..., & Morey, R. D. (this issue). Bayesian inference for psychology, part II: Example applications with JASP. *Psychonomic Bulletin and Review*.
- Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ..., & Morey, R. (this issue). Bayesian inference for psychology, part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin and Review*.
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on *p* values: Context, process, and purpose. *The American Statistician*.