The Journal of Open Source Software

bayes-toolbox: A Python package for Bayesian statistics

Hyosub E. Kim[1,2]

1 School of Kinesiology, The University of British Columbia, Canada **2** Department of Physical Therapy, University of Delaware, United States

DOI: 10.21105/joss.05526

Software

- Review
- Repository
- Archive

Editor: Samuel Forbes

Reviewers:

- @alstat
- @ChristopherLucas

Submitted: 22 April 2023
Published: 27 October 2023

Summary

bayes-toolbox is a Python package intended to facilitate the increased use and adoption of Bayesian statistics in scientific research. As Python is one of the fastest growing and most widely used programming languages, bayes-toolbox fills the need for a Python library that makes it as easy to perform Bayesian statistical tests as it currently is to perform their "frequentist" counterparts. The intended users of bayes-toolbox are students and researchers, particularly those in the behavioral and neural sciences, who are looking for a low friction way to learn Bayesian statistics and incorporate it into their research.

Statement of need

Currently, Python users can choose between several packages that provide simple-to-use functions for running classical/frequentist statistical tests (e.g., Pingouin, SciPy, pandas, and statsmodels). In contrast, for Bayesian statistics there has only been the excellent Bambi package, which, while quite powerful and robust, does require more advanced knowledge and familiarity with R-brms syntax. Therefore, the goal of bayes-toolbox is to fill an important gap in the Python-Bayesian community, by providing an easy-to-use module for less experienced users that makes it as simple, in Python, to fit a Bayesian model to data as it is to run a frequentist statistical test. As all of the models (tests) are executable with single functions, they are ideal for use in an open, replicable workflow (Wilson et al., 2017).

bayes-toolbox is a Python package that makes performing such Bayesian analyses simple and straight forward. By leveraging PyMC, a probabilistic programming library written in Python (Patil et al., 2010), and providing easy-to-use functions, bayes-toolbox removes many of the technical barriers previously associated with Bayesian analyses, especially for users who would prefer to work with Python over other programming languages (e.g., R). This package also removes the requirement to include model formulas to perform statistical tests, another potential barrier for end users. And as the bayes-toolbox functions provide Bayesian analogues of many of the most common classical tests used by scientists, including t-tests, ANOVAs, and regression models, as well as hierarchical (multi-level) models and meta-analyses, it provides a much needed bridge for researchers who are familiar with frequentist statistics but wish to explore the Bayesian framework.

bayes-toolbox was designed for and targeted to researchers primarily in the behavioral and neural sciences. However, as many of the models and Jupyter notebook tutorials included in the public bayes-toolbox repository are adapted from the well-known textbook "Doing Bayesian Data Analysis" (Kruschke, 2014), bayes-toolbox can also serve as an important pedagogical tool for students and researchers alike.

Acknowledgements

Thank you to the PyMC developers, John Kruschke, and Jordi Warmenhoven for generously sharing your work and knowledge.

References

Kruschke, J. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press. https://doi.org/10.1016/c2012-0-00477-2

Patil, A., Huard, D., & Fonnesbeck, C. J. (2010). PyMC: Bayesian Stochastic Modelling in Python. *Journal of Statistical Software*, *35*(4), 1. https://doi.org/10.18637/jss.v035.i04

Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L., & Teal, T. K. (2017). Good enough practices in scientific computing. *PLoS Computational Biology*, *13*(6), e1005510. https://doi.org/10.1371/journal.pcbi.1005510