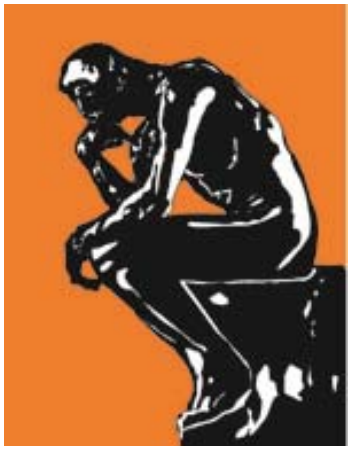
BayesVL Package for Bayesian Statistical Analyses in R

Quan-Hoang Vuong [1,2]
Viet-Phuong La [1,2]

[1] AI for Social Data Lab (AISDL), Vuong & Associates
[2] Centre for Interdisciplinary Social Research, Phenikaa University



Hanoi, April 17, 2019

This short piece serves as a short introduction to BayesVL. BayesVL is our project to develop an easy-to-use, graphics-enabled and pedagogical software package for employing Bayesian statistical analysis for non-expert users in the R statistical programming environment [1].

This R package helps its users to visually learn the graphical structures of Bayesian networks and perform desired simulation exercises together with various technical validations for checking the validity of their models.

The main functions and features of BayesVL Package include:

* Creating the (starting) graphical structure of Bayesian networks;

* Creating one or more random Bayesian networks learned from a dataset with customized constraints;

* Generating JAGS/STAN code for structures of Bayesian networks for sampling and parameter learning;

* Plotting the network graphs and Markov chain Monte Carlo (MCMC) simulations;

* Being compatible with R 2.x or any newer versions.

The development progress of the package is provided in [2], following which the BayesVL Package will ultimately be able to perform such statistical analyses as:

a) Linear regressions for discrete and continuous variables (including interaction terms), following [3,4,5];

b) Bayesian networks, following [6];

c) Nested non-linear hierarchical data models, function well in R, employing both JAGS and Stan MCMC implementations [4,5];

d) GUI for choosing desired models and visual aids like graphs provided in numerous examples of [3,4,5,6], together with the capacity of adjusting models' parameters for practical uses; and,

e) Ready-to-use and user-friendly tools for connecting to datasets/database, especially online publicly deposited ones, in such open repositories as Open Science Framework, Mendeley Data, or Harvard Dataverse, reflecting our ideas of reducing costs of doing science [7].

Inside our GitHub, there will be real-world examples of published results and datasets, which enable BayesVL users to re-run the code and check the outcomes of the simulation exercises by themselves.

Developing BayesVL Package is a continuous process, which requires constant improvements. We hope that finally, the software package will be of use to researchers in the humanities and social sciences who seek to explore the power of Bayesian statistics in application problems [8] and make use of its philosophical values as demonstrated in [3,4,5].

We invite you to check out further progress of BayesVL Package, available at https://github.com/sshpa/bayesvl.

References

[1] Vuong QH, La VP. (2019). BayesVL package for Bayesian statistical analyses in R. *Github*: BayesVL package version 0.6. Available from: < https://github.com/sshpa/bayesvl >

[2] Vuong QH, La VP. (2019). BayesVL: Rules for Versioning. *Open Science Framework*; DOI:10.31219/osf.io/cmuvs. Available from: < https://osf.io/cmuvs >

[3] Kruschke J. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*, 2nd ed. London: Academic Press.

[4] McElreath R. (2018). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. London: Chapman and Hall/CRC.

[5] Vuong QH, La VP, Vuong TT, Nguyen VH, Ho MT, Nguyen THK, Bui QK, Ho MT. (2018). Cultural additivity: Behavioural insights from the interaction of Confucianism, Buddhism, and Taoism in folktales. *Palgrave Communications*, 4, 143. DOI: 10.1057/s41599-018-0189-2. URL: https://www.nature.com/articles/s41599-018-0189-2.

[6] Vuong QH, Bui QK, La VP, Vuong TT, Ho TM, Nguyen HKT, Nguyen NH, Nghiem PKC, Ho MT. (2019, January 26). Cultural evolution in Vietnam's early 20th century: a Bayesian networks analysis of Franco-Chinese house designs. *arXiv Preprints*, arXiv:1903.00817v1 [Stat.AP]; Available from: < https://arxiv.org/abs/1903.00817 > (accessed: April 6, 2019).

[7] Vuong QH. (2018). The (ir)rational consideration of the cost of science in transition economies. *Nature Human Behaviour*, 2(1), 5; DOI: 10.1038/s41562-017-0281-4. Available from: < https://www.nature.com/articles/s41562-017-0281-4 >

[8] Lynch SM. (2007). Introduction to Applied Bayesian Statistics and Estimation for Social Scientists. New York: Springer.