BIOGRAPHICAL SKETCH

NAME: Gabry, Jonah S.

eRA COMMONS USER NAME: NA

POSITION TITLE: Staff Associate (research in Statistics)

EDUCATION/TRAINING

| INSTITUTION AND LOCATION | DEGREE | Completion Date | FIELD OF STUDY |
| --- | --- | --- | --- |
| Bowdoin College (Brunswick, ME) | B.A. | 05/2007 | History, Spanish, Music |
| University of Pennsylvania (Philadelphia, PA) | PostBac | 05/2013 | Mathematics and Statistics |
| Columbia University (New York, NY) | M.A. | 05/2015 | Quantitative Methods in the Social Sciences |

# A. Personal Statement

# A large part of my work focuses on creating methods and tools to be used by applied statisticians. The software packages I develop (described in greater detail in Section C) are intended both to make the process of Bayesian data analysis easier for researchers and to expose them to the methods being investigated by Professor Andrew Gelman and others at the Applied Statistics Center at Columbia University. We are also working on complimentary software tools focusing on multilevel regression and post-stratification of survey data.

# In addition to my work on software for Bayesian data analysis, I am also responsible for survey weights and imputation of missing data for the New York City Longitudinal Survey of Wellbeing, a study carried out by the [Columbia Population Research Center](http://cupop.columbia.edu/research/research-areas/new-york-city-longitudinal-survey-wellbeing-nyclsw).

I look forward to applying my experience in statistical software development and knowledge of hierarchical (multilevel) modeling to the aim of making advanced Bayesian multilevel modeling accessible to clinical data scientists.

# B. Positions and Honors

## Positions and Employment

|  |  |
| --- | --- |
| 2008-2010 | Directory of Customer Accounts, Midwest European Publications, Cambridge, MA & Skokie IL |
| 2010-2013 | Freelance translator and computer programmer, Buenos Aires, Argentina & Philadelphia, PA |
| 2013-2014  2014-2015  2015-present | Graduate research assistant, Columbia University, New York, NY  Earth Institute (Millennium Villages Project)  Graduate research assistant, Columbia University, New York, NY  Political Science Department  Researcher in Statistics (Staff Associate level), Columbia University, New York, NY |

## Other Experience

|  |  |
| --- | --- |
| 2015-Present | Stan Core Development Team |

## Honors and Awards

Quantitative Methods in the Social Sciences Scholarship, Columbia University, 2013

Honors for thesis on Spanish-English translation, Bowdoin College, 2007

Philip C. Bradley Prize, Bowdoin College, 2007

Latin American Studies Prize, Bowdoin College, 2007

Phi Beta Kappa & Cum Laude, Bowdoin College, 2007

Bowdoin College Research Grant, 2006

# C. Contribution to Science

**Co-authored software packages:**

**loo** Efficient approximate leave-one-out cross-validation (LOO) using Pareto smoothed importance sampling (PSIS), a new procedure for regularizing importance weights. As a byproduct of the calculations, we also obtain approximate standard errors for estimated predictive errors, and for the comparison of predictive errors between models. <https://cran.r-project.org/web/packages/loo/index.html>

**rstan** User-facing R functions are provided by this package to parse, compile, test, estimate, and analyze Stan models by accessing the header-only Stan library provided by the 'StanHeaders' package. The Stan project develops a probabilistic programming language that implements full Bayesian statistical inference via Markov Chain Monte Carlo, rough Bayesian inference via variational approximation, and (optionally penalized) maximum likelihood estimation via optimization. In all three cases, automatic differentiation is used to quickly and accurately evaluate gradients without burdening the user with the need to derive the partial derivatives. <https://cran.r-project.org/web/packages/rstan/index.html>

**rstanarm** Estimates pre-compiled regression models using the 'rstan' package, which provides the R interface to the Stan C++ library for Bayesian estimation. Users specify models via the customary R syntax with a formula and data.frame plus some additional arguments for priors. <https://cran.r-project.org/web/packages/rstanarm/index.html>

**shinystan** A graphical user interface for interactive Markov chain Monte Carlo (MCMC) diagnostics and plots and tables helpful for analyzing a posterior sample. The interface is powered by RStudio's Shiny web application framework and works with the output of MCMC programs written in any programming language (and has extended functionality for Stan models fit using the rstan package and the No-U-Turn sampler). <https://cran.r-project.org/web/packages/shinystan/index.html>

**Co-authored papers and conference abstracts/presentations:**

Vehtari, A., Gelman, A., Gabry, J. (2016). *Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC*. Preprint available on arXiv (<http://arxiv.org/abs/1507.04544>)

Andreae MH, White R, Gabry J., Hall C. *Utilization of antiemetic medication as a marker of health care disparities in anesthesia*. Abstract at the Translational Science Conference, Washington, 2015.

# D. Research Support

Ongoing research support

NSF Grant: SES-1424962

PI: Justin Phillips

Award Period: 9/15/14 – 8/31/17

This research focuses on using multilevel regression and post-stratification (MRP) to measure and study dynamic public opinion. My primary responsibilities are to develop MRP methods that extend to time series data and implement software tools using R and Stan to make MRP more accessible to social science researchers working with survey data and facilitate future research on the topic.

INSERT CPRC GRANT INFO HERE

The New York City Longitudinal Survey of Wellbeing is designed to track the dynamics of poverty, hardship, and wellbeing over time among the population of 18+ residents of New York City. My responsibilities include producing imputations of missing data and constructing survey weights to adjust for known differences between sample and population as well as survey non-response.