# Gemma E. Moran


**Contact Information**

425 S. W. Mudd Building, 500 West 120th Street, New York, NY, 10027
Primary email: gm2918@columbia.edu
Secondary email: gem.e.moran@gmail.com
Website: https://www.gemma-moran.com
Phone: +1 (267) 304-4107


**Employment**

**Data Science Institute, Columbia University** — New York, NY, USA

Postdoctoral Research Scientist — August 2019 - Present
Supervisor: David Blei

**Education**

**The Wharton School, University of Pennsylvania** — Philadelphia, PA, USA

Ph.D. Statistics — May 2019
Thesis Title: "Bayesian Approaches for Modeling Variation"
Thesis Advisors: Edward George and Veronika Ročková

**The University of Sydney** — Sydney, NSW, Australia

B.Sc. Advanced Mathematics (First Class Honours) — November 2013
Majors: Mathematics, Statistics
Thesis Advisor: John Ormerod

**Awards**

- Rising Star in Data Science, University of Chicago — 2022
- Poster Award, International Society for Bayesian Analysis — 2022
- Donald S. Murray Prize for excellence in teaching, Statistics Department, Wharton — 2018
- Honorable Mention for Best Talk at the Bayesian Young Statisticians Meeting — 2018
- Wharton Doctoral Program Fellowship — 2014
- The George Allen Scholarship for Mathematical Statistics, the University of Sydney — 2013

**Selected Peer-Reviewed Publications**

* indicates equal contributions

[1] R. Bai*, **G. E. Moran***, J. L. Antonelli*, Y. Chen, and M. R. Boland. "Spike-and-slab group lassos for grouped regression and sparse generalized additive models". In: *Journal of the American Statistical Association* 117.537 (2022), pp. 184–197.

[2] **G. E. Moran**, D. Sridhar, Y. Wang, and D. M. Blei. "Identifiable Deep Generative Models via Sparse Decoding". In: *Transactions on Machine Learning Research*. 2022.

[3] **G. E. Moran**, J. P. Cunningham, and D. M. Blei. "The Posterior Predictive Null". In: *Bayesian Analysis (to appear)* (2022).

[4] **G. E. Moran**, V. Ročková, and E. I. George. "Spike-and-slab lasso biclustering". In: *The Annals of Applied Statistics* 15.1 (2021), pp. 148–173.

[5] **G. E. Moran**, V. Ročková, and E. I. George. "Variance prior forms for high-dimensional Bayesian variable selection". In: *Bayesian Analysis* 14.4 (2019), pp. 1091–1119.

[6] V. Rocková, **G. Moran**, and E. George. "Determinantal regularization for ensemble variable selection". In: *Artificial Intelligence and Statistics* (2016).

## Manuscripts

[7] **G. E. Moran**, D. M. Blei, and R. Ranganath. "Population Predictive Checks". In: *arXiv preprint arXiv:1908.00882 (under review at Journal of the Royal Statistical Society, Series B)* (2022).

[8] **G. E. Moran** and E. I. George. "Nonlinear factor analysis with Bayesian Additive Regression Trees". In preparation. 2022.

## Collaborations

[9] V. Shekar, V. Yu, B. J. Garcia, D. B. Gordon, **G. E. Moran**, D. M. Blei, L. M. Roch, A. García-Durán, M. Ani Najeeb, M. Zeile, P. W. Nega, Z. Li, M. A. Kim, E. M. Chan, A. J. Norquist, S. Friedler, and J. Schrier. "Serendipity based recommender system for perovskites material discovery: balancing exploration and exploitation across multiple models". In: *ChemRxiv* (2022). DOI: 10.26434/chemrxiv-2022-l1wpf.

## Presentations

### Seminars

- "Identifiable variational autoencoders via sparse decoding." Machine Learning Seminar, Microsoft Research New England, December 2021.
- "Identifiable variational autoencoders via sparse decoding." Math and Data Seminar, New York University, October 2021.

### Conferences

- "Identifiable Deep Generative Models via Sparse Decoding." First Workshop on Causal Representation Learning at the Conference on Uncertainty in Artificial Intelligence (UAI), Eindhoven, Netherlands. August 2022. **Invited Speaker**
- "The Posterior Predictive Null." (Poster) International Society for Bayesian Analysis World Meeting, Montreal, Canada. June 2022. **Poster Award (top 10% of posters)**
- "Identifiable deep generative models via sparse decoding." (Invited Talk) International Chinese Statistical Association Applied Statistics Symposium, Gainesville, FL. June 2022.
- "Identifiable variational autoencoders via sparse decoding." (Invited Talk) CMStatistics, London, UK. December 2021.
- "Spike-and-Slab Lasso Biclustering." Joint Statistical Meetings, Vancouver, BC, Canada. August 2018.
- "Spike-and-Slab Lasso Biclustering." Bayesian Young Statisticians Meeting, University of Warwick, UK. July 2018. **Honorable Mention for Best Talk**
- "On variance estimation for Bayesian variable selection." (Poster) International Society for Bayesian Analysis World Meeting, Edinburgh, Scotland. June 2018.
- "Independence Variance Priors for Penalized Likelihood Variable Selection." Joint Statistical Meetings, Baltimore, MA. July 2017.

## Software

- `SSLB` (R Package) **Moran, G. E.** [gemoran/SSLB]
- `SSLASSO` (R Package) Ročková, V. and **Moran, G. E.** [CRAN.R-project.org/package=SSLASSO]
- `EMVS` (R Package) Ročková, V. and **Moran, G. E.** [CRAN.R-project.org/package=EMVS]

## Professional Activities

- Co-Organizer (2016, 2017) Annual Women in Business Academia Conference, Philadelphia, PA

- Board member (2015-2017), Wharton Society for the Advancement of Women in Business Academia
- External Reviewer for *Bayesian Analysis, AISTATS (Selected as a "Top Reviewer" in 2022), NeurIPS, ICML, International Statistical Review, Statistical Methods in Medical Research, Briefings in Bioinformatics*

## Teaching Experience

**Teaching Assistant**

Columbia University

- STCS6701: Probabilistic Models and Machine Learning (Fall 2022)

The Wharton School, University of Pennsylvania

- STAT621: Accelerated Regression Analysis for Business (Fall 2018)
- STAT613: Regression Analysis for Business (Fall 2017, Fall 2018)
- STAT422: Predictive Analytics (Spring 2017)
- STAT101: Introductory Business Statistics (Fall 2016)

**Recitation Instructor**

The Wharton School, University of Pennsylvania

- STAT111: Introductory Statistics (Spring 2018). Rating: 3.1/4

The University of Sydney

- MATH1015: Biostatistics (Semester 1, 2013)
- MATH1005: Statistics (Semester 2, 2013)

**Lecturer**

Wharton Moneyball Academy (Summer 2017, 2018)
*Contributed to and taught course on data analysis in R as part of summer program in statistics and sports for high school students*

## Professional Experience

**Computational Informatics Division, CSIRO** Sydney, NSW, Australia

*Summer research scholar* 11/2013 - 02/2014

Project: Markov chain Monte Carlo and generalized linear models for RNA-seq data

**Australian Mathematical Sciences Institute** Sydney, NSW, Australia

*Summer research scholar* 12/2012 - 02/2013

Project: Theory for Gaussian variational approximation of Bayesian generalized linear models