Katherine Brumberg (Morgan)

kbrum@wharton.upenn.edu | kbrumberg.com

Education

Wharton School of the University of Pennsylvania **Aug. 2019 - May 2024 (Expected)**
PhD in Statistics and Data Science

Yale University **Aug. 2015 - May 2019**
Simultaneous BS/MA in Statistics and Data Science

Research Experience

Graduate Researcher | Advised by Paul Rosenbaum and Dylan Small, Wharton School | Mar. 2020 - Present

- Developed theory about and created an R package for natural stratification, a method which allows for a fixed ratio of controls to treated units in each stratum, enabling the use of strata without needing weights. Controls are selected using randomized rounding of a linear program which quickly returns a feasible solution to the integer program that optimizes covariate balance between the control and treated groups across and within strata. Units may also be simultaneously placed into different comparisons in an optimal way. Applied this work to study the relationship between maternal age and the risk of stillbirth, as well as to the relationship between fluoroquinolone antibiotics and nervous system complications
- Developed theory about and created an R package for splitting propensity score strata into refined strata with improved covariate balance. Strata are decided using randomized rounding of a linear program. Applied this work to a study on right heart catheterization and survival rates

Undergraduate Researcher | Townsend Lab at Yale University | Aug. 2016 - May 2019

- Applied a population genetics and probabilistic framework to cancer evolution research, updating the group's prior model of mutation to account for altered selection due to epistatic interactions

Research Experience for Undergraduates at Mathematical Biosciences Institute | Ohio State University and Harvard School of Public Health | Jun. 2017 - Aug. 2017

- Applied inverse probability weighting to adjust for selection bias in a secondary outcome study
- Used causality model selection techniques to determine ordering of events in prostate cancer study

Books

- Ewens, W., **Brumberg, K**. (2023). Introductory Statistics for Data Analysis. *Springer Nature*. Available from: https://doi.org/10.1007/978-3-031-28189-1.

Publications

- **Brumberg, K.**, Small, D.S. & Rosenbaum, P.R. (2023+). Optimal Refinement of Strata to Balance Covariates. *Submitted.*
- **Brumberg, K.**, Ellis, D.E., Small, D.S., Hennessy, S. & Rosenbaum, P.R. (2023). Using Natural Strata When Examining Unmeasured Biases in an Observational Study of Neurological Side Effects of Antibiotics. *J R Stat Soc Series C*. 72(2):314-329. Available from: https://doi.org/10.1093/jrsssc/qlad010.
- **Brumberg, K.**, Small, D.S. & Rosenbaum, P.R. (2022). Using Randomized Rounding of Linear Programs to Obtain Unweighted Natural Strata that Balance Many Covariates. *J R Stat Soc Series A*. 185(4):1931-1951. Available from: https://doi.org/10.1111/rssa.12848.

- Sinnott, J.A, **Brumberg, K.,** Wilson, K.M., Ebot, E.M., Giovannucci, E.L., Mucci, L.A. & Rider, J.R (2018). Differential Gene Expression in Prostate Tissue According to Ejaculation Frequency. *European Urology.* 74(5):545-548. Available from: https://doi.org/10.1016/j.eururo.2018.05.006.

Software packages

- **Brumberg, K.** (2022). optrefine: Optimally Refine Strata. R package available from: https://CRAN.R-project.org/package=optrefine.
- **Brumberg, K.** (2021). natstrat: Obtain Unweighted Natural Strata that Balance Many Covariates. R package available from: https://CRAN.R-project.org/package=natstrat.

Presentations

- Joint Statistical Meetings (Aug. 2023). "Optimal Refinement of Strata to Balance Covariates." Oral presentation.
- Atlantic Causal Inference Conference (May 2023). "Optimal Refinement of Strata to Balance Covariates." Oral presentation.
- Senior PhD Student Presentations (May 2023). "Optimal Stratification for Covariate Balance." Oral presentation.
- Lawrence D. Brown PhD Workshop (Nov. 2022). "Optimal Refinement of Strata to Balance Covariates." Oral presentation.
- Senior PhD Student Presentations (May 2022). "Optimal Stratification for Covariate Balance." Oral presentation.
- Junior PhD Student Presentations (Sep. 2021). "Neurological Side Effects of Antibiotics: Using Natural Strata to Examine Unmeasured Biases." Poster presentation.
- Joint Statistical Meetings (Aug. 2021). "Obtaining Unweighted Natural Strata that Balance Many Covariates Using Randomized Rounding of Linear Programs." Contributed speed presentation.
- Junior PhD Student Presentations (Apr. 2021). "Using Randomized Rounding of Linear Programs to Obtain Unweighted Natural Strata That Balance Many Covariates." Oral presentation.

Teaching Experience

Professor | Wharton Statistics and Data Science Department | May 2022 - Jun. 2022, May 2023 - Jun. 2023

- Taught the condensed six week summer course "Introductory Statistics" online twice

Teaching Assistant | Wharton Statistics and Data Science Department | Aug. 2021 - May 2022

- Led recitation sections, assisted the professor with course preparation, and graded for the course "Introductory Statistics"

Undergraduate Learning Assistant | Yale Statistics and Data Science Department | Sep. 2017 - May 2019

- Helped instruct students and grade for the courses "Computational Tools for Data Science," "Theory of Statistics," "Data Mining and Machine Learning," and "Data Analysis"

Work Experience

Modeling and Informatics Intern | Vertex Pharmaceuticals | May 2018 - Aug. 2018, May 2019 - Aug. 2019

- Implemented permutation tests to be run alongside the predictive models for compound activity to make sure that released models are performing better than random
- Implemented iterative focused screening to select the best compounds to include in high throughput screens
- Created an RShiny application to detect anomalies in the performance of control compounds over time in an assay

Research Assistant | Yale Statistics and Data Science Department | May 2018 - Aug. 2018

- Developed an RShiny application to help the department match undergraduate learning assistants with courses

Service

Mentor | Wharton Pre-Doctoral Directed Reading Program | Sep. 2022 - Present

- Developed a reading plan with an undergraduate student and met weekly to discuss the reading

Advisory Board Member | Wharton Pre-Doctoral Directed Reading Program | Sep. 2021 - Present

- Matched undergraduate students with graduate members for the program and helped to develop the new program

President | Wharton Society for the Advancement of Women in Business Academia (WSAWBA) | Sep. 2019 - Present

- Led the organization which aims to support women in the Wharton doctoral programs

Pen-pal | Letters to a Pre-scientist | Sep. 2019 - Present

- Wrote letters to a middle school student at a disadvantaged school to broaden awareness of what STEM professionals look like and serve as a source of inspiration

Committee Member | Wharton Statistics and Data Science Quinquennial Review | Jan. 2022 - Apr. 2022

- Gathered doctoral student feedback for the quinquennial departmental review

Mentor | Graduate School Mentoring Initiative | Jan. 2021 - Jun. 2021

- Mentored a first generation low income undergraduate student interested in applying to graduate school

Co-President | Splash at Yale | Sep. 2015 - May 2019

- Led the nonprofit educational outreach organization which brings 1000 middle and high school students to campus to take hour long seminars taught by Yale students

Awards and Scholarships

- 2022 Brown Best Student Paper Award (Wharton School)
- 2022 Donald S. Murray Teaching Prize (Wharton School)
- 2019 George James Doctoral Fellowship (Wharton School)
- 2019 National Science Foundation Graduate Research Fellowship
- 2019 Statistics and Data Science Senior Award (Yale University)
- 2019 Y-Work Award for Outstanding Student Employees (Yale University)
- 2016 Henry S. McNeil Summer Fellowship (Yale University)
- 2015, 2016, 2018 Yale Club of Boston Scholarship (Yale University)
- 2015 Lori Laitman Rosemblum Scholarship (Yale University)

Skills

- Proficient in Linux, Microsoft Office, LaTeX, R, RShiny, Python, SPSS, MATLAB, Mathematica
- Native proficiency in English and Russian, full professional proficiency in French, and proficiency in Latin