Anna Guo

(Updated January 3, 2024)

guoanna.com / Email / Google Scholar / Linkedin / (203)-683-8952

Research Goal: Develop statistical methods that yield robust causal conclusions in the face of mounting data challenges, including missing data, unmeasured confounding, high-dimensionality, and beyond, by leveraging insights from causal graphical models and semiparametric statistics.

Education

Emory University — Atlanta, GA, USA
Ph.D. in Biostatistics (anticipated). Advisor: Razieh Nabi — *Aug. 2021 – May. 2026*

Yale Univeristy — New Haven, CT, USA
Master of Science in Biostatistics — *Aug. 2019 – May. 2021*

University of Wisconsin Madison — Madison, WI, USA
Visiting Student Program in Mathematics — *Sep. 2017 – Dec. 2017*

Zhejiang Normal University — Jinhua, ZJ, China
Bachelor of Science in Mathematics — *Sep. 2015 – June. 2019*

Qualifications

- *Programming Languages:* R, Python, SQL, SAS, Stata, MATLAB, C, Linux, HTML
- *Coursework:* Advanced causal inference, Data mining, Machine learning, Nonparametric statistics, Bayesian statistics, Survival analysis, Statistical inference, Advanced statistical computing, Generalized linear models, Linear regression.

Publications

1. **Anna Guo**, David Benkeser, Razieh Nabi, "Targeted Machine Learning for Average Causal Effect Estimation Using the Front-Door Functional." *arXiv.*
2. **Anna Guo**, Jiwei Zhao, Razieh Nabi, "Sufficient Identification Conditions and Semiparametric Estimation under Missing Not at Random Mechanisms." *Proceedings of the 39th Conference on Uncertainty in Artificial Intelligence (UAI)*, 2023
3. Emma Zang, Poh Lin Tan, Thomas Lyttelton, **Anna Guo**, "Impacts of the COVID-19 Lockdown on Gender Inequalities in Time Spent on Paid and Unpaid Work in Singapore." *Population and Development Review*, 2023
   *Featured (selected): The Straits Times (Singapore)
4. Emma Zang, **Anna Guo**, Christina Pao, Nancy Lu, Bei Wu, Terri R Fried, "Trajectories of General Health Status and Depressive Symptoms Among Persons with Cognitive Impairment in the United States." *Journal of Aging and Health*, 2022

Software

- `fdtmle`: An R package for average causal effect estimation via the front-door functional.

Research Experience

Causal inference methodologies

Causal graphical approach towards missing data

- Presented compelling counterexamples proving that missing not at random (MNAR) mechanism leads to model non-identification;
- Proposed a comprehensive theoretical framework for quantifying and locating additional conditions required for model identification;
- Developed semiparametric estimation strategy that yield unbiased estimation under the challenge of MNAR data.

Targeted machine learning for average causal effect estimation using the front-door functional

- Formulated flexible targeted minimum loss based estimation (TMLE) algorithms tailored for front-door models, capable of handing all types of mediators;

- Adopted super learning techniques to integrate multiple machine learning models, achieving high accuracy in complex model estimation;
- Established the robustness properties of the TMLE estimators via the formulation of the second-order error term.

Application of statistics in social sciences and healthcare

Analysis of longitudinal health trajectories in patients with cognitive impairment

- Extracted and organized longitudinal data concerning individuals with cognitive impairment from the National Health and Aging Trends Survey (NHATS) data using Stata;
- Employed group-based trajectory models to uncover latent groups on health conditions;
- Performed survival analysis to reveal that latent group assignments can accurately predict patients' likelihood of survival or institutionalization.

Analysis of gender inequality

- Utilized fixed effects models to evaluate shifts in spouses' allocation of time to formal and informal works before and after the onset of the COVID-19 pandemic;
- Employed multinomial logistic regressions to examine the influence of demographic and socioeconomic factors on shifts in time allocation across different categories of work.

Intership Experience

Knowbox Technology Co., Ltd. Beijing, China

Data Analyst Internship *Jul, 2018 - Aug, 2018*

- Converted survey data acquired from 300,000 schools into a manageable format using Excel functions;
- Conducted a comparative analysis between survey data and reference data to identify discrepancies and refresh outdated records.

Talks and posters

Sufficient Identification Conditions and Semiparametric Estimation under Missing Not at Random Mechanisms

- Joint Statistical Meeting 2023 *Aug 6, 2023*
- Conference on Uncertainty in Artificial Intelligence *Aug 2, 2023*
- The 2023 American Causal Inference Conference *May 24, 2023*
- European Causal Inference Meeting *April 20, 2023*

Professional Activities

Conference Reviewer

- Conference on Uncertainty in Artificial Intelligence (UAI) *2023*

Teaching Experience

Emory University:

- BIOS-760R: Advanced Causal Inference, Teaching Assistant *Fall 2023*
- BIOS-761: Causal Inference, Teaching Assistant *Spring 2023*
- BIOS-522: Survival Analysis Methods, Teaching Assistant *Fall 2022*

Lanxi No.1 High School:

- Integrated Mathematics II, Intern Math Teacher *Fall 2018*

Honors & Awards

- University of Wisconsin Madison Exchange & Visiting International Student Academic Excellence Award
- Zhejiang Normal University Overseas Study Special Award (With 21500$ scholarship)
- Zhejiang Province Government Scholarship (Top 3% of students within the province)
- Zhejiang Normal University Outstanding Student First Prize Scholarship (Top 5%)