Si (Sylvia) Cheng

sylviascheng.github.io
sylvia-cheng@berkeley.edu

I'm a Ph.D. Candidate extensively trained in causal inference and machine learning. I'm skilled in statistical method development and applications, and experienced in designing ML models for real-world projects using big data.

EDUCATION

University of California, Berkeley, Berkeley, CA 2021-present
Ph.D. Candidate in Epidemiology
M.A. in Biostatistics
Advisors: Alejandro Schuler, Mark van der Laan (D.E. in Computational Precision Health)
- Classes: Theoretical Statistics, Statistical Computing, Causal Inference, Modern Biostatistical Theory and Practice, Survival Analysis and Causality | Teaching: Introduction to Probability and Statistics in Summer 2022 & Summer 2024
- Activity: Student Government Ph.D. Representative, School of Public Health (2022-2023)

University of Rochester, Rochester, NY 2017-2019
M.S. in Epidemiology

Huazhong Agricultural University, Wuhan, China 2013–2017
B.E. in Bioengineering
University of Florida, Visiting Student (2016–2017)

WORK EXPERIENCE

Data Scientist / Biostatistician, Kaiser Permanente Division of Research Mar 2023-Present
- **Individualized Surgical Time Prediction with Uncertainty Quantification**: Designed a meta-learner and built a series of XGBoost models trained on Electronic Health Records to create personalized predictions with uncertainty intervals for surgical duration estimation across 4000+ service categories. Developed a novel conformal quantile regression based method to quantify uncertainty on the individual level with controlled type I error rates. Designed an optimal updating rule for surgical time estimates based on uncertainty distributions, resulting in prediction accuracy comparable to surgeon-level expertise.
- Implemented our decision model to optimize surgical time scheduling across 40+ Kaiser Permanente hospitals to enhance efficiency of perioperative care for 9.4 million KP patients, and presented work at the American Medical Informatics Association Annual Symposium in Nov, 2024.

Data Science Consultant, Apple Inc. (contracted) Oct-Nov 2023
- Designed a data science initiative encompassing data engineering and analysis for a high-impact marketing project targeting a global clientele and audience. Efficiently delivered time-sensitive solutions, enabling the client team to surpass project deadlines while maintaining high-quality standards and client satisfaction.

Graduate Student Researcher, Center of Targeted Machine Learning & Causal Inference (CTML), UC Berkeley Aug 2023-present
- **Perioperative Care Policy Optimization**: Developed the first reinformcement learning model using Conservative Q Learning for severe health outcome prediction and policy optimization for perioperative blood glucose management and treatment at UCSF.
- **High-dimensional Statistical Parameter Optimization for Causal Inference**: Developed a Targeted Maximum Likelihood Estimation (TMLE) based method to identify and optimize high-dimensional longitudinal statistical target parameters describing stochastic interventions, and built a deep learning model to implement the method. Applied the method to estimate the causal effects of an optimized nutrition profile on cardiovascular and aging outcomes of 700 California residents in the NHLBI Growth and Health Study (NGHS), informing future dietary policy changes in California.

Graduate Student Researcher, UC Berkeley Aug 2021-Aug 2023
- **Stochastic Intervention Design & Causal Identification**: Developed machine learning models via Targeted Learning to identify and estimate causal effects of data-adaptive statistical parameters under stochastic interventions of childhood depressive symptoms and cardiovascular health conditions on adulthood aging biomarkers of California participants in NHLBI Growth and Health Study (NGHS).

Research Data Scientist (full-time), University of California, San Francisco Nov 2019-Aug 2021
- Conducted statistical analyses and data engineering for manuscript preparation and publications with professors and postdocs for nation-wide and state-wide research studies of more than 3500 participants, contributed to novel findings in impacts of behavioral and environmental changes on health and mental well-being.

SKILLS

- Programming: R, Python, PyTorch, sklearn, Git, H2O.ai, LaTeX
- Natural Languages: English, Mandarin, Japanese
- Teamwork: Stakeholder communication, project management, public engagement

AWARDS

- Leadership Education in Neurodevelopmental and Related Disabilities (LEND) Fellowship, funded by HRSA's Maternal Child Health Bureau, USA 2018
- National Undergraduate Scholarship awarded by Ministry of Education of China 2016
- Bioengineering Scholarship for Outstanding Undergraduate Students 2015

SELECTED PUBLICATIONS & TALKS

Cheng, S.S., & Schuler, A. (2025). Doubly Robust Policy Learning for Causal Stochastic Interventions through Auto-debiased Neural Networks. Frontiers in Computational Health Conference, Berkeley, CA, United States.

Cheng, S.S., Schuler, A., Cohn, B., & Liu, V.X. (2024). Personalized Uncertainty Quantification in Operating Room (PUQOR): Optimizing Surgical Time Estimation with Conformal Prediction. American Medical Informatics Association Annual Symposium,

San Francisco, CA, United States.

Cheng, S.S., Epel, E., Laraia, B. (2024). Aging Through a Causal Lens: Effects of Depressive Symptoms on Metabolic Dysfunctions & Cellular Aging. UCSF Adolescent & Young Adult Health Research Symposium, San Francisco, CA, United States.

Cheng, S.S., Schuler, A., Epel, E., Laraia, B. (2022, June). Chicken First or Egg First: Examining a Causal Effect of Depressive Symptoms on Metabolic Dysfunctions and Cellular Aging. Society of Epidemiologic Research 2022 Annual Meeting, Chicago, IL, United States.

Rudd, K. L., **Cheng, S. S.**, Cordeiro, A., Coccia, M., Karr, C. J., LeWinn, K. Z., ... & Bush, N. R. (2022). Associations between maternal stressful life events and perceived distress during pregnancy and child mental health at age 4. Research on Child and Adolescent Psychopathology, 1-10.

Blades, R., Mendes, W. B., Don, B. P., Mayer, S. E., Dileo, R., **Cheng, S. S.**, ... & Epel, E. S. (2024). A Randomized Controlled Clinical Trial of a Wim Hof Method Intervention in Women with High Depressive Symptoms. Comprehensive Psychoneuroendocrinology, 100272.