Yuyao Wang

PhD candidate
University of California San Diego
La Jolla, CA 92093 U.S.A.

email: yuw079@ucsd.edu

Education

2019 - now *PhD*, Mathematics with a Specialization in Statistics, University of California San Diego
2019 - 2022 *MA*, Mathematics, University of California San Diego
2015 - 2019 *BS*, Mathematics, Xi'an Jiaotong University

Research Interests

Causal inference, survival analysis, missing data problems, semiparametric theory.

Work Experience

Jun - Aug 2024 Biostatistics internship at St. Jude Children's Research Hospital.

Fellowships

2019 - 2023 UCSD Halicioglu Data Science Institute (HDSI) Graduate Prize Fellowship.

Awards and Scholarships

2024 Society for Causal Inference (SCI) Travel Scholarship (for 2024 ACIC).
2024 UCSD GPSA Travel Grant Award (for 2024 ACIC).
2023 Student Paper Competition Award, 2023 Lifetime Data Science Conference.

Publications and Preprints

2024 **Wang, Y.**, Ying, A., Xu, R. (2024) Doubly robust estimation under covariate-induced dependent left truncation. *Accepted in Biometrika*, asae005, https://doi.org/10.1093/biomet/asae005. *(Won Student Paper Competition Award at 2023 Lifetime Data Science Conference)*

2023 Peng, Y., **Wang, Y.**, Xu, R. (2023). Measures of explained variation under the mixture cure model for survival data. *Statistics in Medicine*, 42(3), 228-245.

Working Papers

2024 **Wang, Y.**, Ying, A., Xu, R. (2024+) Learning treatment effects under covariate dependent left truncation and right censoring. *Working paper.*

2024 White, L. R., **Wang, Y.**, Launer, L., Seale, L., Berry, M., Flannagan, M., Montine, T.J., Xu, R. (2024+) Midlife dietary selenium and Okinawan ancestry are independently associated with less severe Alzheimer neuropathology at autopsy in Japanese-American men. *Submitted to Journal of Alzheimer's Disease.*

Software

R packages: `truncAIPW`, `aftR2`

Professional Service

2024 Student committee member for 2025 International Conference on Health Policy Statistics (ICHPS).

2024 Organizer for the topic-contributed paper session "Causal Inference with Complex Missing Patterns" at 2024 Joint Statistical Meeting (JSM).

Presentations

2024 Learning treatment effects under covariate dependent left truncation and right censoring. *Presentation at 2024 Joint Statistical Meeting (JSM).*

2024 Doubly robust estimation of treatment effects under covariate dependent left truncation and right censoring. *Poster at 2024 American Causal Inference Conference (ACIC).*

2024 Learning Treatment Effects under Covariate Dependent Left Truncation and Right Censoring. *Presentation at Southern California Applied Mathematics Symposium (COCAMS).*

2023 Doubly Robust Estimation under Covariate-induced Dependent Left Truncation. *Presentation at McGill Statistics Seminar (2023 Fall).*

2023 Doubly Robust Estimation under Covariate-induced Dependent Left Truncation. *Presentation*

at 2023 Joint Statistical Meeting (JSM).

2023 Doubly Robust Estimation under Covariate-induced Dependent Left Truncation. *Presentation at 2023 Lifetime Data Science Conference (LiDS).*

2023 Multiply Robust Estimation of Treatment Effect for Time-to-event Outcome under Dependent Left Truncation. *Poster at 2023 American Causal Inference Conference (ACIC).*

2023 Doubly Robust Estimation under Covariate-induced Dependent Left Truncation. *Poster at 2023 Public Health Research Day at UCSD.*

2022 Semiparametric Estimation for Non-randomly Truncated Data. *Poster at 2022 American Causal Inference Conference.*

2022 Semiparametric Estimation for Non-randomly Truncated Data. *Poster at 2022 Public Health Research Day at UCSD.*

Research Experience

2023 - now **Learning treatment effects under covariate dependent left truncation and right censoring**, University of California San Diego, La Jolla, CA, USA.
Student researcher for PhD thesis | Supervisor: Ronghui Xu

- Derived the efficient influence curve for a general Z-estimand under covariate-dependent left truncation.
- Developed a general doubly robust framework for handling covariate-dependent left truncation and right censoring, applicable to a wide range of estimation problems.
- Illustrated the utility of this framework in estimating average treatment effects and heterogeneous treatment effects. For average treatment effects, developed estimators that enjoy model double robustness and rate double robustness; for heterogeneous treatment effects, developed orthogonal and doubly robust learners that can achieve the oracle rate of convergence.
- Carefully studied the asymptotics and convergence rate of the proposed estimators.
- Applied the proposed methods to analyze the effect of midlife alcohol consumption on late-life cognitive impairment, using data from the Honolulu Asia Aging Study.

2020 - 2023 **Selection bias adjustment for time-to-event analysis of prospective childhood cancer survivorship studies**, St. Jude Children's Research Hospital, Memphis, TN, USA.
Biostatistics intern | Supervisor: Kendrick Li

- Developed inverse probability weighting (IPW) approaches to handle selection bias from covariate-dependent left truncation, making full use of the semi-competing risks data.
- Studied the cardiotoxicity of childhood cancer treatments, such as anthracyclines and radiation. Specifically, analyzed the hazard ratio under a Cox model with IPW weights to adjust for left truncation due to delayed entry.

2020 - 2023 **Doubly Robust Estimation under Covariates-induced Dependent Left Truncation**, University of California San Diego, La Jolla, CA, USA.
Student researcher for PhD thesis | Supervisor: Ronghui Xu

- Derived the efficient influence curve (EIC) for the expectation of an arbitrarily transformed survival time.
- Constructed EIC-based estimators that are shown to have favorable properties, including model double robustness, rate double robustness, and semiparametric efficiency.
- Provided technical conditions for the asymptotic properties that appear to not have been carefully examined in the literature for time-to-event data.
- Our work represents the first attempt to construct doubly robust estimators in the presence of left truncation.
- Applied our estimator to analyze data on central nervous system (CNS) lymphoma and data on cognitive impairment from Honolulu Asia Aging study.
- Extended the doubly robust estimating function for other estimands, including causal effects and causal hazard ratio under a marginal structural Cox model.

2023 - 2024 **Studying factors that resist development of Alzheimer's disease neuropathologic changes**, University of California San Diego, La Jolla, CA, USA.
Research Assistant | Supervisors: Ronghui Xu, Lon White

- Data visualization and data analysis for Alzheimer brain autopsy data from Honolulu Asia Aging Study.
- Identified three risk factors (FEV1, dietary selenium, dietary maltose) for Alzheimer brain pathology at death.

2023 **Assessing safety of COVID-19 vaccine exposure during pregnancy**, UC San Diego based Organization of Teratology Information Services (OTIS) research center, San Diego, CA, USA.
Statistical consultant

- Contributed to developing statistical analysis approaches for assessing safety of medications during pregnancy; particularly in addressing challenges in analyzing the effect of different timings of exposure in the presence of selection bias due to left truncation.

2020 - 2022 **Measures of Explained Variation under the Mixture Cure Model**, University of California San Diego, La Jolla, CA, USA.
Student researcher | Supervisor: Ronghui Xu

- Proposed two approaches to define explained variation under mixture cure models: based on the Kullback-Leibler information gain, and based on residual sum of squares.
- Studied properties of the proposed measures both analytically and with simulation studies
- Applied the proposed approaches to analyze data on bone marrow transplant and the SEER-medicare data.

2020 - 2021 **Covariate balancing weights for estimating average treatment effect with censored time-to-event data**, University of California San Diego, La Jolla, CA, USA.
Student researcher | Supervisor: Jelena Bradic

- Reviewed the literature on covariate balancing methods and investigated the possibility of developing covariate balancing weights to estimate average treatment effect for censored time-to-event data.

- Developed method that use balancing weights for confounding and use inverse probability of censoring weighting to handle informative right censoring.

2018 - 2019 **Posterior Consistency for Bayesian Method of Inverse Problems with Non-Gaussian Noise Assumption**, Xi'an Jiaotong University, Xi'an, Shaanxi, China.
Research assistant | Supervisor: Jianxiong Jia

- Studied existing methods and algorithms of Bayesian approach to inverse problems.
- Generalized the consistency result in *Posterior consistency for Bayesian inverse problems through stability and regression results* by Sebastian J Vollmer to Bayesian inverse problem with Gaussian mixture noise.

2018 **Active Subspace and Sliced Inverse Regression**, Georgia Institute of Technology, Atalanta, GA, USA.
Research assistant | Supervisor: Wenjing Liao

- Proved two theorems for the error bounds of the estimated covariance matrix and the estimated projection matrix when using active subspaces.
- Compared the performance of active subspace method with sliced inverse regression in simulation and analyzed convergence rates and find the optimal tuning parameter for active subspace method.

Teaching Experience

University of California San Diego, La Jolla, CA, USA.

2024 MATH 284: Lifetime Data Analysis, TA, Spring 2024.
FMPH 291: Special Topics/Public Health - Applied Survival Analysis, TA, Spring 2024.
DSC 180B: Data Science Project II, TA, Winter 2024.

2023 DSC 180A: Data Science Project I, TA, Fall 2023.
High School Math Program (Probability and Statistics, advanced track), mentor, Summer 2023.
MATH 284: Lifetime Data Analysis, TA, Spring 2023.
MATH 181B: Introduction to Mathematical Statistics, TA, Spring 2023.
DSC 180B: Data Science Project II, TA, Winter 2023.

2022 DSC 180A: Data Science Project I, TA, Fall 2022.
MATH 181A: Introduction to Mathematical Statistics I, TA, Winter and Spring 2022.

2021 MATH 10A: Calculus I, TA, Fall 2021.
MATH 185: Introduction to Computational Statistics, TA, Spring 2021.
MATH 189: Data Analysis and Inference, TA, Winter 2021.

2020 MATH 11: Calculus-Based Probability and Statistics, TA, Fall 2020.
MATH 189: Data Analysis and Inference, TA, Winter and Spring 2020.

2019 MATH 10B: Calculus II, TA, Fall 2019.

Outreach

2023	Presentation at the UCSD Halicioglu Data Science Institute research review event for industry partners
2023	Poster presentation at the UCSD Halicioglu Data Science Institute Open House for prospective PhD students
2022	Moderator for the Prospective International Graduate Student Panel in math department at UC San Diego
2022	Mentor in AWM undergrad mentorship program at UC San Diego
2020 - 2021	Mentor in math graduate mentorship program at UC San Diego

Last updated: August 17, 2024 • Typeset in XeLaTeX