Zhehao (Kenny) Zhang

Email : zhehaoz@uw.edu
Personal Website: kennyzhang-17.github.io
Mobile : +1-805-837-9616

Research Interest: Causal Inference, A/B testing, Sensitivity Analysis, Machine Learning

Education

- **Univeristy of Washington, Seattle** — Seattle, WA
  PhD Candidate in Statistics (with Machine Learning and Big Data track) — *Sept. 2020 – Current*
- **University of California, Santa Barbara** — Santa Barbara, CA
  B.S. in Mathematics; B.S. in Statistics; GPA: 3.94 (Highest Honor) — *Sept. 2016 – June. 2020*

Publication

Z. Zhang, T. S. Richardson. "Bounds on the Distribution of a Sum of Two Random Variables: Revisiting a problem of Kolmogorov with application to Individual Treatment Effects". Link.
MorPhiC Consortium (M. Adli, L. Przybyla, Z. Zhang et al.) "Molecular phenotypes of null alleles in human cells (MorPhiC) consortium: Towards functional characterization of all human genes". Accepted by Nature.
C. Segal, Z. Zhang, B. T. Karras, D. Revere, G. Zane, J. G. Baseman. "Early Epidemiological Evidence of Public Health Value of WA Notify, a Smartphone-based Exposure Notification Tool: Modeling COVID-19 Cases Averted in Washington State". Link.
Z. Zhang, N. LaPierre, B. Hill, C. Cinelli. "PySensemakr: Sensitivity Analysis Tools for Regression Models in Python". Under Review by Journal of Statistical Software. Link.
Z. Zhang, C. Cinelli. "Exact and Approximate Inference for Omitted Variable Bias". In preparation.
Z. Zhang, T. S. Richardson, J. Robins. "Individual Treatment Effect: Prediction Intervals and Sharp Bounds". In preparation.
B. Bhaskar; J. Alumbaugh; Z. Zhang; J. Dillon; A. Burke. "A Comparison of Maxillofacial and Head Injuries Following Electric Scooter and Bicycle Accidents". Link.
S. Panesar; J. Van; Z. Zhang; J. Dillon. "Does a Submental Airway Compared to a Tracheostomy Reduce Length of Stay in Craniomaxillofacial Trauma?". Link.

Industrial Experience

- **Databricks** — San Francisco, CA
  Data scientist intern. Revenue forecasting using iTransformer models and PyTorch. — *Jun.2024 - Sept.2024*
- **Fred Hutchinson Cancer Center** — Seattle, WA
  Part-time research assistant. ATAC-seq data, gene regulatory network, large scale causal discovery. Jan.2024 - Jun.2024
- **Google Cloud Platform** — Sunnyvale, CA
  Software Engineer (PhD) Intern. Deploy G Suite search features. — *Jun.2022 - Sept.2022*
- **Comcast Applied AI Team** — Washington, DC
  Research Intern. Run offline experiments in Link. — *Jun.2021 - Sept .2021*

Research Experience

- **Personalized Decision Making Using Causal Inference** — Washington, WA
  Research Assistant in Department of Statistics — *Sept.2021 - Present*
  - **Research Topics**: Supervised by Professor Thomas Richardson. Characterized prediction intervals for the individual treatment effect (ITE) and established sharp bounds on the distribution functions of ITE. **Preprint will be available on ArXiv soon.**
- **Causal Inference and Sensitivity Analysis** — Washington, WA
  Research Assistant in Department of Statistics — *Sept.2021 - Present*
  - **Research Topics**: Supervised by Professor Carlos Cinelli. Developed new methods to perform hypothesis testing on causal effects under unmeasured confounders.
  - **Others**: Developed sensitivity analysis tool PySensemakr package in Python with over 10,000 downloads. Github.

- **WA Notify Data Analysis and Evaluation Team; University of Washington** — Washington, WA
  Research Assistant with Department of Public Health — *Feb.2021 - Jun.2021*
  - **Modeling and Analysis**: Worked with Washington State Department of Health, Apple and Google to evaluate the effectiveness of Bluetooth notification technology to eliminate the transmission of COVID-19.
  - **Others**: Large-scale data manipulation; Time series analysis; Privacy and fairness of data.
- **Toronto Western Hospital (Fujitsu); University of Toronto** — Toronto, ON
  Summer IMS Researcher (with Dr. Mojgan Hodaie's group) — *Apr.2020 - Sept 2020*
  - **Automatic Brain Tumor Segmentation**: Use a U-net based deep learning architecture to perform image semantic segmentation on brain MRI images with a focus on Trigeminal Neuralgia.
  - **Others**: Gamma Knife optimization algorithm based on sphere packing. Multi-Modality MRI image generation by conditional General Adversarial Networks.
- **Statistics Department, UCSB** — Santa Barbara, CA
  Researcher, Senior Thesis (with Prof. Alex Shkolnik) — *Jan.2019 - Apr.2020*
  - **Optimal James-Stein Shrinkage for Regression**: Develop a new James-Stein type estimator for cross-sectional ordinary least square regression with asymptotic optimization guarantee on dispersion bias. Provide theoretical guarantees and numerical experiments. Thesis.
- **Fields Institute for Mathematical Science** — Toronto, ON
  Summer REU Researcher (with Prof. Andreas Hilfinger) — *June.2019 - Sept.2019*
  - **Inverse Problem for Stochastic Models**: Inferred rate functions for complex stochastic models in biological processes. Used supercomputer to simulate large-scale continuous time Markov Chain and developed algorithms for solving linear network topology models based on Hill functions. Preprint.
  - **Others**: Presented at Pacific Math Alliance Conference and Undergraduate Mathematics Symposium.

Projects

- **Option Trading Visualization App**: Write a Python Dash App for option trading visualization. Link.
- **Recommender Systems Against Shilling Attacks**: Evaluated the Robustness of Collaborative Filtering Recommender Systems against Shilling Attacks. Implemented 4 algorithms and more than 20 attacks. Link.
- **Hull Tactical ERP prediction contest**: Investigated stock return prediction using Long Short Term Memory (LSTM) models. Won most creative category with $1000 in contest. Link.
- **Time Series Analysis (PSTAT 274)**: Built a Time series model to predict on 5-year break-even inflation rate. Link.
- **Statistical Machine Learning (PSTAT 231)**: Used the dataset of 2016 president election and US census to build a model for election prediction. Data organization, visualization. Link.
- **Statistical Learning and Computer Vision (STAT 535)**: Used the CIFAR-10 data set to build a convolutional neural network model for classification (PyTorch). Link.
- **Statistical Learning with Sparsity (STAT 538)**: A detailed study on a Variational Inference paper used for posterior sampling with implementation (Julia, Turing.jl). Link.

Teaching and Service

Service: Reviewer for Journal of Econometrics ..... Mar 2024 — Present
Service: Reviewer for UAI 2024 ..... Feb 2024 — Present
Service: Managing the UAI Mailing List ..... Sep 2021 — Sept 2022
Service: Organizing Committee of Causal Reading Group ..... Sep 2022 — Present
Teaching Assistant: Statistics, UW
Causal Inference Module of Summer Instistute in Statistics and Modeling in Infectious Diseases. ..... Jul 2023
STAT 502 Design and Analysis of Experiments. ..... Jan 2022 — Mar 2022
STAT 504 Applied Regression. ..... Jan 2022/2023 — Mar 2022/2023
STAT 566 Causal Modeling. ..... Mar 2023 — Jun 2023
STAT 570 Advanced Regression. ..... Sept 2023 — Dec 2023
STAT 571 Advanced Regression Methods for Dependent Data. ..... Jan 2024 — Mar 2024
STAT 396 Finite Markov Chains and Monte-Carlo Methods. ..... Mar 2022 — Jun 2022
STAT 311 Elements of Statistical Methods. ..... Sep 2020 — Mar 2021
Directed Reading Program: Statistics, UW
Teach and guide project on A/B testing. ..... Jan 2024 — Mar 2024
Teach and guide project on Deep Learning and Computer Vision. ..... Jan 2021 — Mar 2021
Teach and guide project on Causal Inference. ..... Sep 2021 — Dec 2021

Honor and Awards

REU fellowship at Fields Institute in University of Toronto ..... Jun 2019
Most creative category in Hull Tactical ERP Prediction Contest ..... May 2019
Putnam Mathematical Competition, top 5 in UCSB ..... Dec 2017
College of Creative Studies Honor ..... Jan 2017 — Mar 2020
Dean's Honor in College of Letters and Science ..... Sept 2016 — Mar 2020
American Math Competition top 5% worldwide ..... Mar 2016

Coursework

Probability Theory (A), Stochastic Calculus (A), Statistical Learning (A), Time Series (A), Matrix Analysis (A), Regression Methods (A), Advanced Theory for Statistical Inference (A-), Causal Inference (A), Real Analysis (A), Linear Algebra (A+), Stochastic Process (A+), Probability Theory (A+), Bayesian Analysis (A).

Programming Skills and Research

- **Languages**: Python, R, Julia, C++, LaTeX **Research**: Causal Inference, Regression, Machine Learning