Elvis (Han) Cui

213-317-7388 | elviscuihan@g.ucla.edu | Los Angeles, CA | Website: https://elviscuihan.github.io/

EDUCATION

University of California, Los Angeles   California, United States
- PhD in Biostatistics (GPA: 3.91/4.0)   Expected June, 2024
- PhD Topic: Applying Metaheuristics to solve complex optimization problems with applications to bioinformatics, computer vision and construction of various types of optimal design for biomedical problems with minimal assumptions.

University of California, Los Angeles   California, United States
- MS in Biostatistics (GPA: 3.95/4.0)   June, 2020
- Master Report: Markovian Structures in Modelling COVID-19

Zhejiang University   Hangzhou, China
- BS in Agriculture and Environmental Science (GPA: 3.95/4.0)   June, 2018
- Thesis: Plant Nutrition Deficiency Detection Using Convolutional Neural Networks

AWARD AND HONORS

- **Irma Polaski 2-year Fellowship** for a foreign student in the Fielding School of Public Health for academic excellence.

TECHNICAL SKILLS

- **Languages:** Python, R, SAS, MATLAB, WinBUGS, Stata, JAVA, C, SQL, Julia.
- **Courses (already taken or planning to take before summer 2022):** Core classes in Biostatistics, Optimal design with applications to biomedical studies, survival analysis, longitudinal data analysis, spatial data modeling, adaptive designs for clinical trials, Markov Chain Monte Carlo, Graphical models, linear statistical models, multivariate analysis, stochastic models, and sampling methods.
- **Teaching Assistant** for the following classes to date: Biostat 250A (Linear Models), Biostat 203A (SAS and R), Biostat 203B (Data Science) and Biostat 100A (Stata).

PUBLICATION AND PRE-PRINT

1. Collins, M. D., **Cui, E. H.**, Hyun, S. W and Wong, W. K. (2021). A Model-based Approach to Designing Developmental Toxicology Experiments using Sea Urchin Embryos. Archives of Toxicology. In press. (https://elviscuihan.shinyapps.io/Dc_optimal_design/ for R Shiny App).

2. **Elvis Han Cui**, Dongyuan Song, Weng Kee Wong and Jessica Jingyi Li (2021), "Single-cell generalized trend model (scGTM): a flexible and interpretable model of gene expression trend along cell pseudotime". Submitted to Bioinformatics. Under review. (https://www.biorxiv.org/content/10.1101/2021.11.25.470059v1).

3. Yanan Li, Pengyang Li, **Elvis Han Cui**, Donghui Wang (2021), "Inference Fusion with Associative Semantics for Unseen Object Detection". Proceedings of the AAAI Conference on Artificial Intelligence. 2021, 35(3): 1993-2001. (https://ojs.aaai.org/index.php/AAAI/article/view/16295).

4. **Elvis Han Cui**, Bingbin Li, Yanan Li, Weng Kee Wong, Donghui Wang (2021), "Trajectory-aware Principal Manifold Framework for Few-Shot Image Generation". Submitted to Computer Vision and Pattern Recognition (CVPR).

5. Bingbin Li, **Elvis Han Cui**, Yanan Li, Weng Kee Wong, Donghui Wang (2021), "Dual Path Structural Contrastive Embeddings for Learning Novel Objects". (https://arxiv.org/pdf/2112.12359).In preparation.

6. **Elvis Cui** and Heather Zhou (2019), "Projection pursuit with applications to scRNA sequencing data". (https://arxiv.org/abs/1912.07602). In preparation.

7. **Elvis Han Cui** and Weng Kee Wong (2022), Charaterization of D-Optimal Designs for Two-Parameter Binary Regression Models with Various Link Functions. In preparation.

8. **Elvis Han Cui**, Weng Kee Wong and Minjeong Jeon (2022), Metaheuristic algorithms with applications to maximum likelihood estimation in item response models. In preparation.

RESEARCH EXPERIENCE

Department of Biostatistics, UCLA   Supervisor: Gang Li

Graduate Student Researcher   Dec 2021 - Present

- Developing semi-parametric Bayesian methods and algorithms for instrumental variable analysis with applications to analyzing the UK Biobank (UKB) cancer data. I worked with two-stage regression models with right and interval censored data, Cox's multiplicative model, AFT model and competing risk models.

- Analyzed Women's Healthy Eating and Living (WHEL) data using different survival regression models stratified for various covariates as risk factors. Performed exploratory data analysis using R (KMsurv, survivor, ggplot2 packages).

Department of Statistics, UCLA   Supervisor: Jessica Jingyi Li

Graduate Student Researcher   March 2021 – November 2021

- Initiated and led a group to develop computational and visualization tools for scRNA data. Developed scGTM, a Python package for building kinetics models of two variables with various shapes of correlation (https://github.com/ElvisCuiHan/scGTM) (*See publication and pre-print 2 and 6*).

- Identified immune genes with increasing patterns using scGTM and gene oncology (GO) analysis. Such findings help biologists to detect genes that of great medical and biological interest. Shortened ~50% running time on three large genomic (scRNA seq) datasets by proposing a modified optimization algorithm (PSO) to enhance large scale and parallel computation.

School of Education and Information Studies, UCLA   Supervisor: Minjeong Jeon

Graduate Student Researcher   March 2020 – July 2020

- Designed various EM algorithms for large scale data mining with x1000 times faster than SOTA packages in R and Julia (https://github.com/ElvisCuiHan/computational-methods-for-large-scale-data-mining) (*See publication and pre-print 8*).

- Helped to improve the quality of Math teaching in elementary schools by analyzing longitudinal educational data using item response theory (IRT) models.

Department of Biostatistics, UCLA   Supervisor: Weng Kee Wong

Finding optimal designs in a two parameter binary regression model   April 2021 –July 2021

- Created a web-based tool for implementing a R-code on the Shiny Platform to find different types of optimal designs for a toxicological study using sea urchins, and developed a characterization theorem based on an equivalence theorem to confirm optimality of a 2-point design for regression models with a binary outcome. (See publication and pre-print 1 and 7, respectively).

WORK EXPERIENCE

Artificial Intelligence Research Center, Zhejiang Lab   Hangzhou, China

Artificial Intelligence Researcher   Sept 2020 – Aug 2021

(This work was done under confinement in China due to the pandemic.)

- Currently have published/submitted three manuscripts on top conferences in artificial intelligence and machine learning (*See publication and pre-print 3, 4 and 5*). Research projects has fostered a collaboration with an AI research company to improve the quality and efficiency of data augmentation in aerospace industry.

- Designed a manifold learning framework for data augmentation and few shot image classification problem, which improved 30% accuracy of on the benchmark few shot datasets and enhanced the smoothness of interpolation by 50% of the baseline on benchmark datasets. Developed a fast projection embedding algorithm for few shot object detection models and shortened ~10% running time of the models (Faster-RCNN, Yolo-V3 and other few shot object detection models).

Sperax, I*nformation Technology and Services*   Menlo Park, California, Work Remotely

Statistician   Aug 2020 – Oct 2020

- Developed a stochastic Markov dynamic model framework to simulate the stream of nodes of the system in Python.

- Results helped the team to make decision on the distribution of Sperax coin and are used to give a better understanding of the working mechanism of Sperax coin.