Jeremiah Zhe Liu

Department of Biostatistics
Harvard University

zhl112@mail.harvard.edu
jereliu.github.io
github.com/jereliu
+1 (319) 594-4694

Education

Harvard University (Boston, MA) *PhD Biostatistics, Minor in Computer Science* Expected 2018
Completed advanced doctoral curriculum. GPA: 3.94/4.00.

Research concentrated on Kernel-based prediction/testing under Dr. Brent Coull.

University of Iowa (Iowa City, IA) *BS Statistics, Mathematics, Computer Science* May 2013
magna cum laude, GPA: 3.96/4.00.

Technical Skills

- **Analysis & Modelling:** R, Matlab, Python, SAS, Stan, BUGS, Mplus
- **Graphics & Documents:** ggplot2, OpenGL, C (GTS), Shiny, ArcGIS, LaTeX
- **High Performance Computing:** C (OpenMP, CUDA, OpenCL)
- **Software Development:** C++, Java, Python, Shell script

Professional Experience

Harvard Clean Air Research Center 2013-Present.
Assistant Statistician

- Building prediction system for heavy-metal air pollutants by integrating information from various data sources (air monitoring records, satellite images, etc) under Random Forrest and Kernel Regression.
- Developed Python and R script to extract 3D GIS features from openstreetmap API. Implemented cluster-end Python program to restructure high-volumn wind-trajectory data.
- Implemented automated feature selection for GIS features using a combination of measurement error-based weighting and Ridge-type penalization. Conducted stratified cross validation to assess the model's out-of-sample prediction and the influence of prediction error on the risk estimation in second-stage association studies.
- Developing new methods to conduct semi-supervised clustering of spatial regions to ehance prediction accuracy.

Research Experience

Testing for Interaction between Correlated High-dimensional Kernel Effects 2014-Present.
Advisor: Dr. Brent Coull & Dr. Xihong Lin

- Proposed novel hypthesis test for the interaction effect under Tikhonov-regularized Gaussian Kernel Machines.
- Developing strategy for efficient computation of test statistic's distribution.
- Plan to expand the test to the interactions between more general kernel effect specifications.

Causal Networks for Retarded Bone Growth in HIV-infected Adolescents 2014-2015.
Advisor: Dr. Brent Coull, Dr. Jane Lindsey & Dr. Denise Jacobson

- Trained regularized probabilistic network to model the association between biomarkers and bone growth measures in prenatally HIV-infected adolescents in PACTG 1045 study.
- Wrote interface between M-plus and R to deploy computation-intensive hypothesis generation and testing.
- Identified sub-collection of biomarkers robustly associated with retarded bone growth in HIV-affected teenagers.

Open Source Project

GURLS_MKL: Fast Multiple Kernel Learning Library for GURLS Package 2015

- Independently developed multiple kernel learning functionality for *Grand Unified Regularized Least Squares* (GURLS), an state-of-art supervised-learning package developed at MIT
- Extended fast Proximal Forward-Backward Splitting (PFBS) optimization algorithm to allow memory-efficient iteration update with parallel support. Derived boundary conditions on algorithm parameters to guarantee model convergence.

GPU-Accelerated Sampling for Bayesian Normal Conditional Autoregressive Models 2012
PI: Dr. Kate Cowles

- Designed and implemented parallel algorithms in OpenCL for new model computation strategy proposed by Cowles et al.(2012) for Bayesian Normal CAR model.
- Implementation incorporated into R package *CARrampsOcl*.

Coursework

Statistical Machine Learning

- Kernel Method Theory • Advanced Regression & Learning
- Semi-parametric Inference • Computation Intensive Statistics

Statistics

- Probability Theory • Advanced Bayesian Inference • Theory of Hierarchical Linear Models
- Environmental & Spatial Statistics • Analysis of Genetic Association Studys
- Causal Methods for Mediation and Interaction

Mathematics

- Linear Algebra & Multivariate Calculus • Real Analysis & Measure Theory • Matrix Theory
- Numeric Analysis • Ordinary Differential Equation • Nonlinear Optimization

Computer Science

- High Performance & Parallel Computing • Data Structure • Algorith • Foundation in GIS

Publications

Liu Z, Lindsey J, Coull B, Jacobson D. *Biomarkers and bone growth across Tanner stages in perinatally HIV- exposed youth in PACTG 1045. In Progress*

Zhang J, Gao J, **Liu Z**, et al. *Short-term Effects of the 2013 Beijing Haze Episode on Local Hospital Outpatient and Emergency Room Visits*. Journal of Environment and Health. *To appear*

Liu Z, Zhang J, Zhao B, et al. *Population-based reference for birth weight for gestational age in northern China*. Early Human Development 2014;90(4):177-87.

Honors & Other Activities

HSPH Central Grant, Department of Biostatistics, Harvard School of Public Health, 2014-2015

Department Scholarship, Department of Statistics & Act. Sci., University of Iowa, 2013

ISEE Conference Student Scholarship, International Society of Environmental Epidemiology, 2012-2014

Annual Departmental Scholarship, Department of Statistics & Act. Sci., University of Iowa, 2012

Phi Beta Kappa, Alpha of Iowa Chapter, CLAS, University of Iowa, 2012