Sibei Liu

sl4660@columbia.edu (646)361-2079 www.linkedin.com/in/sibei-liu-biosta

Education

Columbia University, Mailman School of Public Health, New York, NY *May,2021(Expected)*

Master of Science in Biostatistics (Theory and Method Track) GPA:3.75

Related Courses: Probability, Biostatistical Methods (regression modeling, survival analysis, longitudinal analysis), Data Science 1, Data Science for Machine Learning (Regularization, non-linear regression, Classification), Statistical Inference, Epidemiology, Advanced Statistical Computing (Monte Carlo, Optimization Algorithms, Bootstrap, MCMC), SAS (Observer)

Jilin University, School of Public Health, China *June,2019*

Major: Information Management & Information System

Honors :**05/2018** China Scholarship Council- Mitacs (Canada) Summer Internship funding

09/2017 China National Scholarship in Academic Year 2016-17 awarded by Ministry of Education of China (0.2%)

In-Class Academic Projects

Columbia University, *New York, NY*

Bayesian Model for hurricane speed *May, 2020*

- Designed an MCMC algorithm to estimate the posterior distribution of model parameters
- Estimated parameter using posterior mean and 95% credibility interval and evaluated the prediction performance

Daily COVID-19 Cases Prediction model Across Nations *April, 2020*

- Estimated the coefficients of logistic curve model by using the Adam algorithm and visualized data
- Used Mixed Gaussian Model performed by EM algorithm to cluster the estimated parameters of each nation

Optimization on breast cancer data *March, 2020*

- Use Newton-Raphson to estimate of the coefficients logistic regression model
- Built a logistic-Lasso model to select features and used Path-wise Coordinate-wise optimization to find a path of solutions with a sequence of lambdas
- Used Cross-validation to select the best lambda with highest AUC

Simulation study to compare three survival models *February, 2020*

- Used Monte Carlo algorithm to generate 3 sets of survival data with fixed true parameters which follow Exponential proportional hazard model, Weibull proportional hazard model and Cox proportional hazard model respectively
- Fitted three models under each set of data and simulated it for 10000 time a set of betas were extracted and used to calculate the mean and 95% CI to do the comparison with true parameter

Examine Gender Pay Gap at Houston College of Medicine *December, 2019*

- Used R to visualize data and built an association model
- Quantified associations between salaries and gender with other covariates included by performing multiple linear regression

The influencing factors of U.S. Suicide Rate *November, 2019*

- Extracted data from 11 website, produced world suicide graphics by using R
- Performed multiple linear regression to find the influencing factors of suicide
- Built a github-based website to show the project report (https://ys3298.github.io/life_saver.github.io/index.html)

Research Experience

Columbia University, New York, NY

Risk Factors of Myelitis Encephalitis based on three hospitals *Research Assistant* *May,2020 - current*

- Outputted the descriptive tables of social and clinical covariates
- Imputed the missing data and did logistic regressions to explore the risk variables.

Meta-Analysis projects *Data Analyst* *November, 2019-January, 2020*

- Used random effects model to find the pooled mean differences or ORs and produced forest plots in R
- Examined sources of between-study heterogeneity by using subgroup analysis in R
- Mastered network meta-analysis method in R

McGill University, Montreal, Canada

Child Maltreatment and Genetic Variation in Adult Psychopathology *Research Assistant July, 2018- October,2018*

- Extracted available data from publications and completed quality evaluation for screened papers
- Divided and summarized quantitative and qualitative papers
- Wrote the manuscript and submitted to Journal of Affective Disorders

Jilin University, Changchun, China

The Influencing Factor of Teenager Smoking in Liaoning Province *Research Assistant March, 2018-May,2018*

- Determined the innovative perspective of data analysis
- Used the complex sampling pack in SPSS to analyze multistage stratified cluster sampling, including completing Chi-square analysis and the logistic analysis to find risk factors of teenager smoking

Publications

- Pallavi Misra, ***Sibei Liu***, Xiangfei Meng, What DNA methylation modifications and/or genetic variations interact with childhood maltreatment in the development of depression: A systematic review, Journal of Affective 252,2019, 325-333
- Tingting Gao et al, including ***Sibei Liu***, The Role of School Connectedness and Maladaptive Cognitions in the Association between Stress and Internet Addiction: A Serial Mediation Model, Perspect Psychiatr Care. 2019; 55: 728‐733
- Muzi Li, ***Sibei Liu***, Xiangfei Meng, Interactions of child maltreatment and genetic variations in adult psychopathology: A systematic review, Journal of Affective Disorders (Submitted)

Skills

Good command of R, SPSS, Microsoft Office (Access, Excel, Word, PowerPoints).
Entry level of SAS
Languages: English, Mandarin (Native)