Qingdou (Paulina) Han

Paulinahan15@gmail.com | 607-601-0176 | New York, NY

EDUCATION

Columbia University, School of Public Health, New York, NY *Expected May 2023*
Master of Science in Biostatistics

Beijing Normal University (BNU), School of Statistics, Beijing, China *July 2021*
Bachelor of Science in Statistics

- GPA: 87/100

University of California, Berkeley, Berkley, CA *May 2020*

- Exchange program, Overall GPA: 4.0/4.0.

SKILLS

Language: Proficient in English and Chinese (Test Result: TOEFL: 107 GRE: 327 (AW: 3.5))
Computer: Intermediate C, Proficient in R, sql, Python, SAS, Eviews

RESEARCH EXPERIENCE

Estimation of the infarct volume in patients with ischemic stroke *Aug.2020-Mar.2021*
Advisor: Professor Gaorong Li

- Performed univariate analysis on high dimensional clinical data.
- Employed inverse probability of treatment weighting method based on propensity score to fill in missing response values.
- Conducted SCAD penalty to realize variable selection in linear model.
- Compared the MSE between model without implementation and model without variable selection.

Use of Statistical Method among BNU Undergraduate Students *Mar.2018-Mar.2019*
Advisor: Professor Jiao Jin

- Collected the research paper of undergraduate students across 22 departments from websites by web scraping.
- Cleaned and classified the data by the use of different statistical methods
- Applied cluster analysis and ANOVA analysis among different clusters and compared their preference of the usage of statistical methods
- Visualized the outcome and concluded the characteristics of different departments based on their preference in statistical methods.

INTERNSHIP EXPERIENCE

Deloitte Touche Tohmatsu Certified Public Accountants LLP Beijing Branch, Data Analyst Intern
Nov.2020-Feb.2021

- Studied the given business process and determined the relevant sources of information under supervision.
- Conducted tests both traditional and specially adapted tailored to the needs of different client together with team members.
- Performed data analysis and visualization of results.

Chinese National Center for Cardiovascular Diseases, Department of Medical Statistics, Beijing
Intern *Aug.2020-Oct.2020*

- Studied FDA guidelines of follow-on companion diagnosis and evaluation of drug efficiency.
- Drew report of the statistical evaluation of FCD via the external concordance study and survival analysis using Cox model and Bayesian model.

RELEVANT PROJECT

Time series analysis on COVID-19 data *(Introduction to Time Series)* *SPRING 2020*

- Transformed data into growth rate of the COVID-19 cases and logarithm form.
- Employed ARIMA models and deterministic models on the transformed truncated dataset.

- Used Rolling cross validation and cross validation for model selection and made predictions.

Study on the Main Factors Affecting Divorce Based on Statistical Methods and Gottman Partner Treatment *(Statistical Learning)* *FALL 2019*

- Verified the effectiveness of Gottman Partner Treatment by applying KNN and SVM classification methods.
- Employed logistics regression on the training data set.
- Analyzed the main factors of divorce by employing PCA, Lasso and Random Forrest on the training data set and compared their results on the testing data set.

Comparison between different methods of variable selecting *(Linear Model)* *FALL 2019*

- Coded LQA adaptive Lasso and Lasso, AIC, BIC variable selecting guidelines by R.
- Employed LQA adaptive Lasso, Lasso, stepwise and AIC, BIC guidelines methods on the simulated data set.
- Tested the efficiency of different methods.

Factors of Power Consumption of Residential Electrical Appliances based on Multi-Statistical Methods *(Applied Multivariate Statistical Analysis)* *FALL 2019*

- Used k-means Cluster Analysis to divide the records of power consumption into 2 groups based on the different environmental conditions.
- Reduced the dimension of dependent variables using PCA and employed multivariate linear regression.
- Utilized Stepwise, Lasso, GBM, to select important variables, and draw conclusion of the relationship between some specific environmental factors and Power consumption.

Study on City Satisfaction based on Weibo Comment *(Mathematical Model)* *FALL 2018*

- Cleaned over 30000 pieces of Weibo Comments by python using regular expression and SVM.
- Graded every comment under different topics with NLP method.
- Proposed a city's satisfaction statistic specifically designed for Beijing government using the entropy weight method.
- Conducted robust test on the model.

Analysis of Economic Influencing Factors of Beijing-Tianjin-Hebei Coordinated Development *(Econometrics)* *FALL 2018*

- Collected relevant data from Website and cleaned the raw data by completing the missing data using Fourier Fitting Employed logistics regression on the training data set.
- Constructed Logarithmic model on the data
- Conducted Q-test, ARCH test, Ramsey RSET test, redundant variables test and correlation test on the model.

EXTRACURRICULAR ACTIVITIES

Member, **Workshop of Public Health in Tsinghua University** *Jun.2020-Aug.2020*
Vice President, **Dancing Troupe of Beijing Normal University** *Sep.2018 –July.2019*
Member, **Public Relation Department of Student Association of BNU** *Sep.2017–July.2018*