Chen Li

(475)-243-4567
chenlimichelle@hotmail.com
100 Haven Ave Tower 3 20E, New York, NY, 10032

Education

Columbia University Mailman School of Public Health May 2019
- M.S. – Biostatistics (GPA 4.08/4.00)
- Related coursework: Survival analysis, Topics in statistical learning & data mining, Design of medical experiments, Pharmaceutical statistics

Peking University Health Science Center July 2017
- M.S. – Pharmaceutical Analysis (GPA 3.67/4.00)
- B.S. – Pharmacy (GPA 3.22/4.00)

Skills

- Proficient in using R and SAS for statistical analysis, simulation, modeling, and data manipulation
- Proficient in Microsoft Office(Excel, Outlook, Word, Powerpoint, Access)
- Familiar with MySQL and relational database management system
- Language: English, Mandarin, Conversational German

Experience

Boehringer Ingelheim Pharmaceuticals Inc. June 2018 - Aug. 2018
Summer Intern, Biostatisics and Data Science
- Developed R Shiny app for exploratory analysis and visualization of Analysis Data Model(ADaM) datasets
- Applied cluster analysis, dimension reduction and differential expression analysis to RNA-seq data
- Generated analysis datasets according to CDISC Study Data Tabulation Model (SDTM)

Columbia University Mailman School of Public Health Sept. 2018 - Dec. 2018
Teaching Assistant, Statistical computing with SAS
- Held office hours to help graduate students from various backgrounds understand statistical analysis
- Gave brief talks in class about key concepts and practices in homework

China National Health Development Research Center Apr. 2016 - July 2016
Research Intern
- Set up a database on hospital construction status by merging data from 1000+ hospitals using Excel
- Literature review on healthcare system integration

Course Projects

Prediction Modeling for Happiness Score Apr. 2018
- Applied statistical learning methods(ridge regression, lasso, PCR, PLSR, KNN, SVM) to build prediction models for happiness scores using information from 9 variables
- Used resampling methods to compare the distribution of prediction errors of regression models
- Used ensemble tree-based methods(bagging, boosting, random forest) to build prediction models and identify important variables in prediction

Data Visualization on Leading Causes of Death Nov. 2017
- Developed R Flexdashboard and R Shiny to create dynamic visualization of public health data
- Used ggplot2 and tidyverse to visualize and summarize mortality rate among regions through 10 years

Clinical Trial Data Analysis Feb. 2018

- Analyzed Pharmacokinetics data from a phase I clinical study to test for bioequivalence
- Performed a dose-response analysis using multiple comparison procedure
- Analyzed data from a rheumatoid arthritis study using Cochran-Mantel-Haenszel test and generalized linear mixed effect model

Certification

SAS Base programming in SAS 9

Publication

Wu, Ruijun, Chen Li, Cong Li, Jinyu Ren, Xiaozhi Sun, Sufang Zhang, Juncheng Zou, and Xiaomei Ling. "Rapid Screening of Multi-target Antitumor Drugs by Nonimmobilized Tumor cells/tissues Capillary Electrophoresis." Analytica Chimica Acta (2018).