Hong Xu

+86 18851671537 hongxulalala@163.com HongXu.github,io

EDUCATION

Minzu University of China **Sep. 2023 – Jun. 2025**
School of Science, Master of Applied Statistics. GPA: 3.79/4 *Beijing, China*

Nanjing University of Finance and Economics **Sep. 2019 – Jun. 2023**
School of Applied Mathematics, Bachelor of Financial Mathematics. *Nanjing, China*

RESEARCH EXPERIENCE

Well-Calibrated Risk Prediction Model for Binary Outcomes with Missing Data **Jul.-Oct.2023**

- Assisting in developing predictive models for breast cancer risk by proposing a constrained calibration model to address discordance between target and source populations using two-phase data.
- Conducting numerical simulations to assess model robustness across various distributions and scenarios with R.

Risk Estimation and Evaluation Using Semiparametric Models in Two-Phase Data **Dec.23-Mar.2024**

- Utilizing a semiparametric model to estimate the predicted risk and the AUC under the Two-Phase data. The model evaluation was conducted through simulations based on various two-stage sampling methods, including case-control, simple random sampling, balanced sampling, and R-balanced sampling.

Survival Analysis of Adverse Outcomes in STEMI Patients Based on Troponin Levels **Jul.-Aug.2024**

- Developing survival analysis and Cox multivariate regression,evaluated the role of TPS levels in risk stratification. Assessing the impact of these novel predictors by calculating the AUC, IDI,NRI within the Cox model.(Manuscript Submitted)

Multi-Task Learning for Disease Prediction Using Polygenic Risk Scores **Sep.-Nov.2024**

- Developing a ML model to predict disease susceptibility with PRS in a multi-task learning framework. Employed resampling techniques to address data imbalance and utilized a MLT with an attention mechanism to extract both general and task-specific features.

Assessing the Added Value of New Predictors in Cardiovascular Disease Prevention **Dissertation**

- Evaluating the improvement in predictive models by incorporating novel risk factors.Demonstrated that evaluation metrics (IDI, NRI, AUC) are U-statistics with specific statistical properties.
- Developing semi-parametric statistical methods to compute AUC, IDI, and NRI in the missing data(Two phase data), updating the R package *"TwoPhaseAccuracy"* to include these methodologies.

PROFESSIONAL EXPERIENCE

Vanke School of Public Health, Tsinghua University, RA **Apr.-Nov.2024**

- Exploring the discrimination faced by adolescence pregnant girls in SSA based on the DHS database, and the mediating effect of discrimination on adverse premature births and current health outcomes.

Centre for Quantitative Medicine, Duke-NUS Medical School, RA **Sept.2024-present**

- Conducting genetic analysis on polypoidal choroidal vasculopathy (PCV) and age-related macular degeneration (AMD) in Singapore Chinese case-control cohorts.To integrate data from Japanese and Korean cases and extracted overlapping variants between public control datasets and our own cohorts.

Health Inequality Among Middle-aged and Elderly People in China **Mar.-Jun.2024**

- Studied health inequalities in China using the Charles 2020 database; employed the LightGBM model for feature selection and developed an Ordinal Regression Model with CLMM and Empirical Bayes estimation to estimate mean and variance of health levels across provinces.

Clinical Intelligent Diagnostic Modeling for Hemorrhagic Stroke **Sep.-Oct.2023**

- Processed Big Data in EHRs using machine learning, applied the Boruta algorithm to identify key predictors for hematoma expansion and used XGBoost regression to predict 90-day mRS scores from extensive feature sets.

PUBLICATION & MANUSCRIPT

- YQ Cao*, **H Xu**. Two-phase Data Analysis Methods and Application[J]. *Journal of Minzu University of China(Natural Sciences Edition)* (Accepted).
- HJ Wang, **H Xu**, JW Wang *. Associations of the TyG-Related Indicators and Overactive Bladder: Enhancing Risk Stratification in Middle-Aged and Elderly Populations[J].*International Journal of Surgery*(Under review).
- Jia L*, **Xu H**, Xu X, et al. GmMYB114 Facilitates the Synthesis of Anthocyanins in Soybean Sprouts under Blue Light[J].*Plants*, 2024,13(8):1107.
- YQ Cao*, **H Xu**. Development and evaluation of the updated risk prediction model involving in new candidate predictors(Manuscript completed)

TECHNICAL SKILLS

Languages: Mandarin(native), English(proficient), Cantonese(proficient)
Skills:R, Python, SQL, PS, ACCESS, Eviews, SPSS, Microsoft Office, LaTeX