Yichen He

E-mail: e0732876@u.nus.edu *Telephone number:* +65 89429892
Place of birth: Nanjing, Jiangsu, China *Date of birth:* 10-02-2000

Education

Ph.D's degree in Biostatistics *National University of Singapore*
Year 1 *August 2024 - Now*

- Research Interests : Infectious disease dynamic modelling and cancer progression natural history modelling

Master's degree in Statistics and Data Science *National University of Singapore*
CAP: 4.5/5.0 *August 2021 - December 2023*

- Courses : Data Modelling and Computation, Discrete Optimization, Numerical Methods in Differential Equations, Stochastic Process, Categorical Data Analysis, Statistical Foundations of Data Science, Applied Regression Analysis, Deep learning and applications.
- Capstone research project: Distribution-free Multisample Test Based on Optimal Matching with Applications to Feature Selection

Bachelor's degree in Mathematics *Soochow University*
GPA: 3.8/4.0 *September 2018 - July 2022*

- Courses : Calculus, Advanced Algebra, Ordinary Differential Equations, Probability, Complex Analysis, Abstract Algebra, Numerical Analysis, Topology, Number Theory, Measure and Integration, Numerical Analysis.
- Capstone project: A Comparative Study of Image Recognition Algorithms based on Feature Extraction

Work experience

Research Assistant *April 2023 - August 2024*
Saw Swee Hock School of Public Health, National University of Singapore *Singapore*

- Disease modelling and Statistical analyses.
- Academic writing and publication of results.
- Preparation of meeting materials for stakeholders.

Research experiences

Demographic Epidemiological Modelling Of Singapore (DEMOS) *April 2023 - August 2024*
Research Assistant (ongoing)——Team Member

- Utilize mathematical techniques to dynamically update demographic modules, including mortality, fertility, BMI, and other relevant factors respect to different combinations of age, gender and race group.
- Construct a lung cancer incidence model based on comprehensive screening data from the population of Singapore.
- Apply rigorous statistical analysis on the entire model and dataset to ensure accuracy and reliability of the results.
- Enhance the model by incorporating socio-economic information, such as housing prices and education level, to better quantify individuals' susceptibility to various diseases and assess their vulnerability.

Distribution-free Multisample Test Based on Optimal Matching with Applications to Feature Selection *February 2023 - June 2023*

Graduate research project (ongoing)—-Team Leader

- In this paper, we apply a new graph-based multi-sample distribution-free test which utilizes non-bipartite optimal matching and Mahalanobis distance construct the test statistic A in order to measure the similarity between different samples and provide a view of underlying structure.
- In addition, for the purpose of better capturing pivotal variables in high-dimensional data, we introduce a recently proposed method called Graph-based Feature Selection (GFS) algorithm to perform a hierarchical agglomerative clustering procedure on data.
- By combining the two methods together, we can manage to do experiments on high-dimensional dataset under general alternatives

Examining the Dose–Response Relationship between Outdoor Jogging and Physical Health of Youths: A Long-Term Experimental Study in Campus Green Space *June 2022*

Publications (Int. J. Environ. Res. Public Health 2022, 19(9), 5648; —-Co-author)

- Based on physical health tests for 2852 youths from a Chinese university from September 2018 and September 2019, using China's National Student Physical Health Standard (NSPHS).
- Applying statistical testing like paired t-tests to examine statistical differences and using multiple regression model to evaluate the associations between jogging and physical health.
- The Paper analyze the main factors affecting physical health of Chinese students, classifies the statistical relationships between different environmental factors and give implications for improving health.

Technical skills

Programming Languages	C, C++, C#, Python, R, MATLAB, SQL
Tools	LaTeX, MS office

Language proficiencies

Chinese	Native
English	Fluent in communication and writing / IELT 7.0 / GRE 159+170+3.5