

Kang Shuai
Geisel School of Medicine, Dartmouth College
The Dartmouth Institute for Health Policy & Clinical Practice
Kang.Shuai@dartmouth.edu
<https://shuai-kang.github.io/>

Education

2016 - 2020 B.S., Mathematics and Applied Mathematics, Zhejiang University, Hangzhou, China

2020 - 2025 (expected) Ph.D., Statistics, Peking University, Beijing, China
Advisor: Yangbo He

Visiting Experience

Jan. 2024 - June. 2024 Visiting Student Researcher, Statistics, UC Berkeley, California, United States
Advisor: Peng Ding

Research Interests

Economics, Instrumental variable, Factorial experiment, Health-related data analysis, Mediation analysis, Non-Gaussian data, Peer effect, Sequential treatments.

Publications

Kang Shuai, Shanshan Luo, Yue Zhang, Feng Xie, Yangbo He. “Identifiability and estimation of causal effects with non-Gaussianity and auxiliary covariates”, to appear in *Statistica Sinica* (40 pages), 2024.

Kang Shuai, Shanshan Luo, Wei Li, Yangbo He. “Identifying causal effects using instrumental variables from the auxiliary dataset”, to appear in *Statistica Sinica* (39 pages), 2024.

Working Paper

Kang Shuai, Lan Liu, Yangbo He, Wei Li. “Mediation analysis with unmeasured confounding between parallel mediators and outcome”, submitted to *Electronic Journal of Statistics*.

Shanshan Luo*, **Kang Shuai*** (Co-first author), Yechi Zhang, Wei Li, Yangbo He. “Identification and estimation of causal peer effect using instrumental variables”, submitted to *Journal of Business & Econometric Statistics*.

Data Analysis Experience

- We apply our method to a real-world dataset investigating the impact of the trade on income using cross-sectional data for the year 2019 from the World Bank and use 154 records for analysis. We finally demonstrate our method can yield similar conclusions without requiring a valid instrumental variable. This is from **the first paper of my publication**.
- We use a primary dataset (7539 samples) from the National Health and Nutrition Examination Survey (NHANES) program during 2011-2012, and another auxiliary dataset (2571 samples) from the population-based study Monica10. We analyze the causal effect of Vitamin D deficiency on BMI in the primary dataset by using an instrumental variable (a binary filaggrin genotype) from the auxiliary dataset. We show that vitamin D does not have significant effects on BMI, which is consistent with many previous findings. This is from **the second paper of my publication**.
- We analyze a mouse obesity dataset (287 samples) in order to investigate the mediation role of gene expressions from a single nucleotide polymorphisms (SNPs, exposure) to the body weight (outcome). We finally identify two active mediators (gene expressions **Slc39a11** and **Ercc3**) linking SNP to body weight. The point is that we consider the presence of unmeasured confounding (some unmeasured phenotypes) between mediators and outcome. This provides insight for studying the mediation mechanism with unmeasured confounders. This is the direct application of our method from **my mediation paper** submitted to **Electronic Journal of Statistics**.
- We apply our methods to data from the China Family Panel Studies (CFPS), conducted by the Institute of Social Science Survey, Peking University, between 2010 and 2018. We will investigate the potential spillover effect of the ego's retirement on the peer's cognitive health by using couples' dataset from CFPS. This study aims to understand the potential effects of retirement on both individuals and their spouses, providing policymakers with insights into retirement systems and health interventions. We ended up with paired samples of 1373 couples, totaling 2746 individuals. Our results appeared to be promising for understanding the potential spillover effect of retirement on cognitive health. This is from **our peer effect paper** submitted to **Journal of Business and Economic Statistics**.

Seminar and Conferences

North-South Six Universities Mathematics Doctoral Forum (First Prize), Wuhan University, Wuhan, China, 2023.

The 7th PKU-Tsinghua Statistics Forum (Best Poster Award), Peking University, Beijing, China, 2023.

The Inaugural Causal Inference Section of the Chinese Association for Applied Statistics (Oral Presentation), Beijing Technology and Business University, Beijing, China, 2023.

The 2nd National Joint Conference on Statistics and Data Science (Oral Presentation), Kunming, China, 2024.

The 9th National Graduate Statistics Forum (Top Ten Paper Award), Renmin University of China, Beijing,

China, 2023.

The American Causal Inference Conference, Seattle Washington, United States, 2024.

Teaching

Teaching Assistant, Advanced Mathematics, Peking University, Fall 2020-Spring 2021 & Fall 2022-Spring 2023.

Teaching Assistant, Linear Algebra, Peking University, Spring 2022-Fall 2022, Spring 2023-Fall 2023 & Fall 2023-Spring 2024.

Teaching Assistant, General Statistics, Peking University, Spring 2021-Fall 2021 & Fall 2021-Spring 2022.

Honors and Scholarship

Outstanding Student Award of Peking University, 2020-2021.

Minlue Innovation Scholarship of Peking University, 2021-2022.

Outstanding Research Award and Jiukun Scholarship, 2022-2023.

Outstanding Research Award, Research Innovation Award, and BYD Scholarship, 2023-2024.

Referee Service

Journal: Biometrika, Statistics in Medicine, Journal of Machine Learning Research.

Conference: International Conference on Learning Representations, International Conference of Machine Learning.

Technical Skills

Programming: R, Python, C, Matlab

Professional Softwares: R, Python

Drawing & Typesetting: \LaTeX , Beamer

Languages: Chinese (Native), English