



清华大学统计学研究中心
Center for Statistical Science, Tsinghua University

年度报告

Annual Report

2022.07-2023.06
July 2022 to June 2023

北京·清华园





CONTENTS

目录

1

中心概况

2

组织架构

- 学术委员会
- 顾问委员会
- 杰出访问教授
- 行政班子
- 行政团队

学科团队

- 中心教员
- 博士后
- 科研助理
- 博士研究生

3

学科发展与人才引进

- 学科发展
- 教员晋升
- 合作研究

4



6

学术活动

- 主办学术活动
- 统计学与数据科学论坛
- 参加学术活动

8

社会服务及影响

- 社会服务
- 学术杂志服务

5

学术成果

- 学术论文
- 专利及软著
- 科研项目
- 奖励荣誉

人才培养

- 本科生培养
- 研究生培养
- 优秀大学生夏令营

7



国际著名统计学家、清华校友、美国哈佛大学刘军教授和林希虹教授一直密切关注中心的发展动态，并从人才培养、团队建设、学科发展等多个方面给予支持和指导。

中心概况

清华大学统计学研究中心是学术独立的校级研究中心，统筹规划清华大学统计学科发展建设，行政上挂靠工业工程系。中心自 2015 年成立以来，始终秉承并践行“开拓创新、争创一流”的发展理念，以“建立高水平师资队伍，开展高水平学术研究，推动跨学科交叉合作，建设国际一流学科”为发展目标，推动学科发展建设。

目前，中心已初步建成一支以优秀青年人才为主的朝气蓬勃的研究团队，并在清华园建立起完整的统计学人才培养体系，涵盖从本科到博士、博士后各个层次。中心现有教授 1 人、杰出访问教授 1 人、客座教授 3 人、副教授 10 人、助理教授 2 人、讲师 1 人、博士研究生 45 人（截至 2023 年 6 月 30 日）。

在学术研究方面，中心以统计学理论和方法研究为基础，着重推动生物健康统计、经济金融统计、工业统计与运筹学、统计机器学习等交叉研究前沿方向，取得了丰硕的学术研究成果。上述多方面努力推动清华大学统计学科快速发展，学术声誉不断提升，国内外影响力持续提高。





- 6月，中心成立仪式
- 9月，首届博士研究生入学



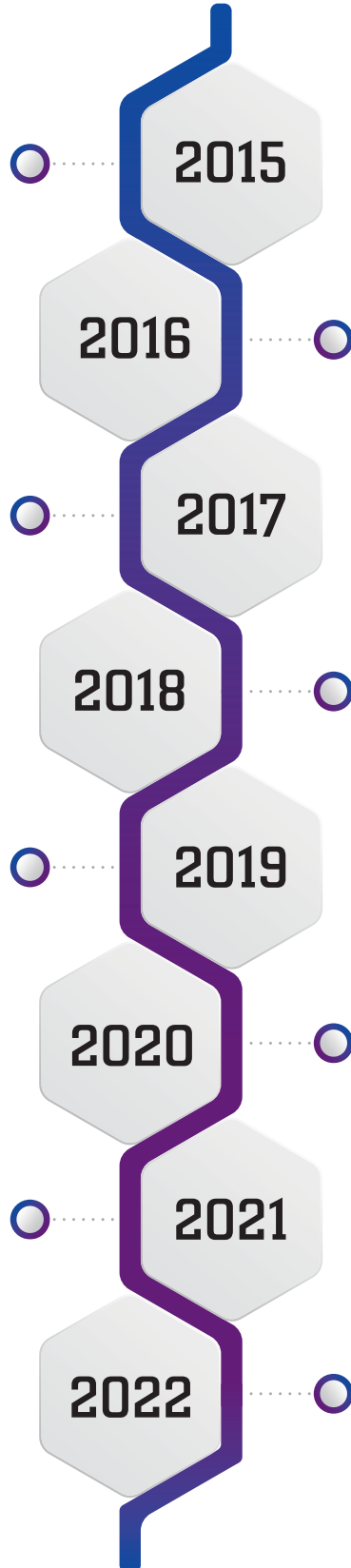
- 9月，开设本科生“统计学辅修”项目



- 11月，承办“国际计算统计协会亚洲分会 25周年大会暨中国现场统计学会计算统计分会第二届年会”
- 7月，“泰晤士高等教育中国学科评级” A+



- 12月，中心师生在“清华大学抗击新冠肺炎疫情表彰大会”荣获表彰



- 3月，发起成立“中国现场统计研究会计算统计分会”
- 5月，建立“清华大学统计咨询中心”
- 9月，入选教育部“双一流”学科建设名单

- 7月，成立“中国统计咨询合作联盟”



- 10月，聘请汤家豪院士出任“杰出访问教授”

- 1月，清华大学“统计学”本科一学位通过教育部备案
- 3月，“统计学与运筹学”学科在QS世界学科排名保持第16名



- 8月，邓柯课题组成果“技术贸易措施综合指数体系”、俞声课题组成果“生物信息学本体系统(BIOS)”受邀出席服贸会

组织架构

学术委员会



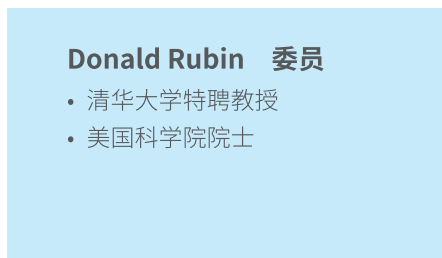
刘军 主任

- 哈佛大学统计系终身教授、生物统计系教授



杨立坚 副主任

- 清华大学统计学研究中心教授



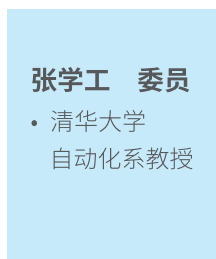
Donald Rubin 委员

- 清华大学特聘教授
- 美国科学院院士



王凯波 委员

- 清华大学万科公共卫生与健康学院教授



张学工 委员

- 清华大学自动化系教授



苏良军 委员

- 清华大学经济管理学院教授



孙茂松 委员

- 清华大学计算机系教授



顾问委员会



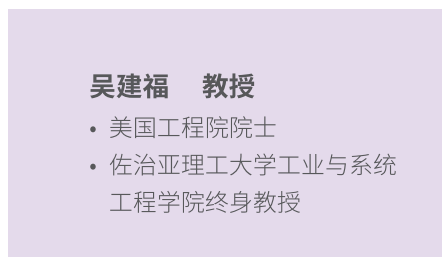
耿直 教授

- 中国概率统计学会前理事长
- 北京大学教授



Donald Rubin 教授

- 清华大学特聘教授
- 美国科学院院士



吴建福 教授

- 美国工程院院士
- 佐治亚理工大学工业与系统工程学院终身教授



王永雄 教授

- 美国科学院院士
- 斯坦福大学统计系终身教授



Terry Speed 教授

- 澳洲科学院院士
- 加州大学伯克利分校统计系和 WEHI 终身教授

杰出访问教授



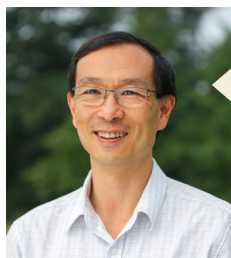
汤家豪 院士

- 伦敦政治经济学院 荣休教授
- 挪威科学与文学院 外籍院士

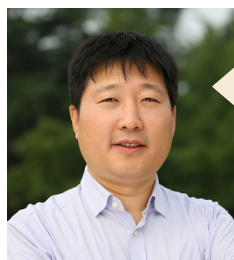
行政班子



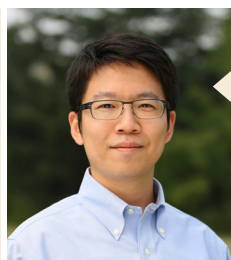
侯琳 副教授
副主任



杨立坚 教授
学科发展主管



李东 副教授
本科生教学主管



俞声 副教授
研究生教学主管



田园
行政与安全主管、
行政秘书

行政团队



田园
办公室主任



侯禹珊
宣传主管



马腾
行政助理



王泽
行政助理

学科团队

- 中心教员 -



杨立坚 教授

- 清华大学统计学研究中心长聘教授
- 北卡罗来纳大学教堂山分校统计学博士
- 美国统计协会会员 (ASA Fellow)
- 国际统计学会当选会员 (ISI Elected Member)
- 国际数理统计学会会士 (IMS Fellow)
- 国际工程技术协会杰出会士 (IETI Distinguished Fellow)
- 国家级人才计划入选者

研究方向: 时间序列、函数型及高维数据的统计推断, 以及统计学在经济学、金融学、农学、食品科学、地理学、遗传学、神经科学和管理科学的应用



邓柯 副教授

- 清华大学统计学研究中心长聘副教授
- 北京大学统计学博士
- 哈佛大学统计系博士后、副研究员
- 北京智源人工智能研究院“智源研究员”
- 国家级青年人才计划入选者

研究方向: 贝叶斯统计、统计计算、生物信息、文本分析、人工智能方法



李东 副教授

- 清华大学统计学研究中心长聘副教授
- 香港科技大学统计学博士
- 香港科技大学数学系博士后
- 爱荷华大学统计与精算系博士后

研究方向: 复杂时间序列的统计分析、非欧数据分析、空间统计、网络数据分析、机器学习、金融计量学

- 清华大学统计学研究中心长聘副教授
- 北京大学统计学博士
- 耶鲁大学生物统计系博士后、副研究员

研究方向: 统计遗传学、生物信息学、应用统计



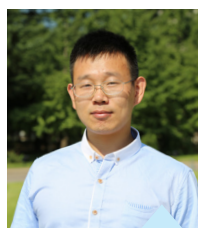
侯琳 副教授

- 清华大学统计学研究中心长聘副教授
- 乔治华盛顿大学系统工程(运筹学)博士
- 哈佛大学医学院博士后
- 国家级青年人才计划入选者

研究方向: 医学统计、自然语言处理、电子病历数据分析、医学知识提取、临床决策支持



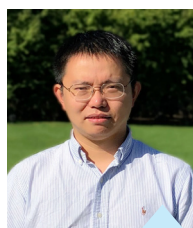
俞声 副教授



刘汉中 副教授

- 北京大学统计学博士
- 加州大学伯克利分校统计系博士后
- 国家级青年人才计划入选者

研究方向: 高维数据统计推断、因果分析



林乾 副教授

- 麻省理工学院数学博士
- 哈佛大学统计系博士后
- 北京智源人工智能研究院“青年科学家”
- 国家级青年人才计划入选者

研究方向: 降维方法、函数型/拓扑型数据分析、蒙特卡洛方法

- 多伦多大学统计学博士
 - 伦敦大学学院统计系博士后
 - 鲁尔波鸿大学数学系博士后
- 研究方向:** 时间序列、变点推断、M估计、网络数据

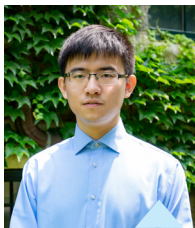


吴未迟 副教授

- 德克萨斯 A&M 大学统计学博士
 - 哥伦比亚大学生物统计系博士后
- 研究方向:** 分位数回归、测量误差分析、高维数据统计分析、流行病学与生物遗传学的统计分析、电子医疗病历数据分析



王天颖 副教授



杨朋昆 助理教授

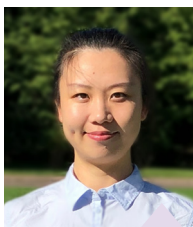
- 伊利诺伊大学香槟分校电子与计算机工程博士
 - 普林斯顿大学电子工程系博士后
 - 国家级青年人才计划入选者
- 研究方向：高维统计理论、机器学习、算法及优化



张静怡 助理教授

- 乔治亚大学统计学博士
- 研究方向：数据融合、数据降维、去中心化网络、最优传输理论

- 北京大学统计学博士
- 宾夕法尼亚大学联合培养博士
- 复旦大学上海数学中心博士后



邓婉璐 教学副教授

- 清华大学统计学博士



周在莹 教学副教授



王江典 讲师

- 北卡罗来纳州立大学统计学博士

- 博士后 -



李艺超

- 合作导师：邓柯
 - 清华大学统计学专业理学博士
 - 2021年7月加入中心
- 研究方向：统计计算、贝叶斯分析



赖建发

- 合作导师：林乾
 - 香港浸会大学统计学专业理学博士
 - 2021年11月加入中心
- 研究方向：神经网络理论研究



郑家森

- 合作导师：林乾
 - 中国人民大学统计学专业理学博士
 - 2021年11月加入中心
- 研究方向：高维数据、神经网络

- 科研助理 -



宋希婷

邓柯课题组



张起帆

侯琳课题组

- 博士研究生 -

2018级: 黄 昆、罗赛迪、罗声旋、潘长在、沈 翀、宋泽宁、王 掣、杨萱铃、余 博、朱 珂

2019级: 任吉杨、宋 爽、孙 爽、陶宇心、王海洋、吴方维、郑思捷、周墨钦

2020级: 白露佳、冯永真、胡祺睿、李冬煜、卢伟灏、卢 鑫、王 达、徐曼芸、余 成、苑洪意、张卓婧

2021级: 付子初、韩庭萱、李弘梓、陆 瑶、罗天派、马 运、王羽超、王梓涵、易盈淮、于丁一、张皓博、赵政昀

2022级: 蔡乐衡、范歆远、江柔蓝、李易诚、罗颖橙、吕逸晨、马昕桐、潘庆一、孙弘毅、应怀原、于浩洋、张灿睿

学科发展与人才引进

- 学科发展 -



汤家豪 院士



汤家豪院士续聘清华大学杰出访问教授

2022年10月, 汤家豪院士正式续聘“清华大学杰出访问教授”, 这是汤先生自2019年10月加盟清华大学统计学科的第二个聘期。自受聘以来, 汤先生在推动清华大学统计学科发展、人才培养、科学研究等方面付出诸多努力与贡献。

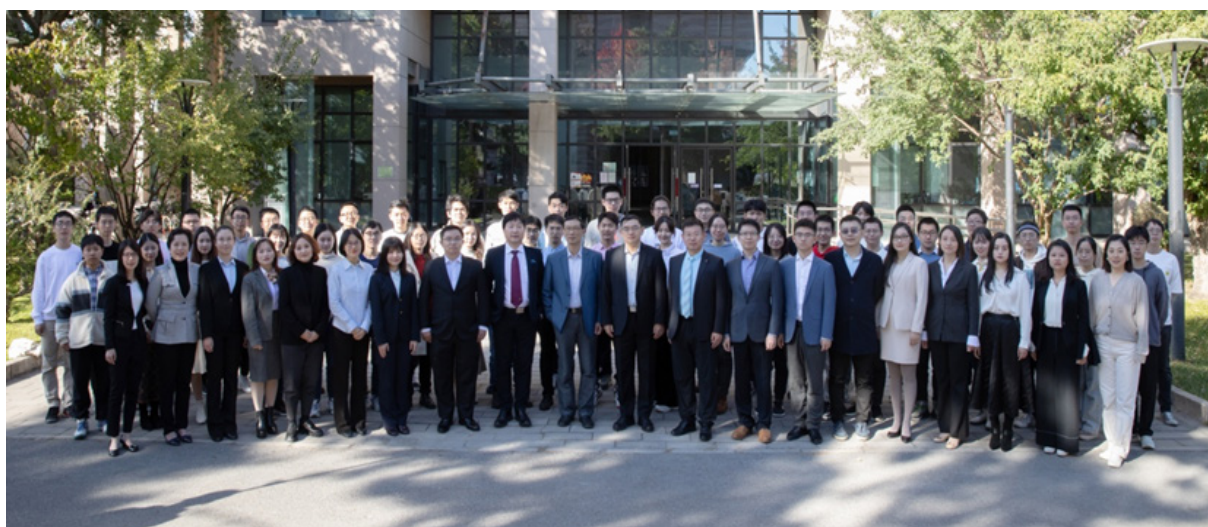
- 教员晋升 -



邓婉璐
晋升
教学系列副教授



王天颖
晋升
准聘副教授





- 合作研究 -

清华大学统计学研究中心高度重视产学研的有效结合。自 2016 年起，先后为原国家质量监督检验检疫总局（现国家市场监督管理总局）、国家食品安全风险评估中心、海关总署、国家市场监督管理总局等多个政府部门的政策决策和改革方案制定提供重要技术支持，为我国进出口食品安全监管改革、食品安全评估不确定性分析、日本输华食品核污染风险评估、新冠疫情防控、食品安全评价性抽检方案制定等工作做出重要贡献。

中心与北京协和医院、北京清华长庚医院、北京大学第一医院、北京大学第三医院、粤港澳大湾区数字经济研究院（福田）等多个医院或研究机构深入合作，运用统计学优势助力智能医学诊断系统、电子病历数据分析、医学知识图谱建设等应用的开发。

此外，中心与国内外知名人文社科机构联合进行了大量交叉研究，运用数据科学手段协助人文历史学者整理海量中国经典历史文献和非物质文化遗产资料，提高研究效率，促进文化传承与创新。



学术成果

- 学术论文 -

- ◆ Jie Li and Lijian Yang* (2023). Statistical inference for functional time series. *Statistica Sinica* 33 (1): 519-549.
- ◆ Kun Huang, Sijie Zheng and Lijian Yang* (2022). Inference for dependent error functional data with application to event-related potentials. *TEST* 31 (4): 1100-1120.
- ◆ Jia Peng*, Youfa Wang, Ming Yang, Limin Wang, Xuchao Yang, Xinyu Shi, Lijian Yang et al (2022). Inequalities of spatial primary healthcare accessibility in China. *Social Science & Medicine* 314: 115458.
- ◆ Shen Chong, Minyue Pei#, Xiaoxiao Wang, Yiming Zhao, Luning Wang, Jiping Tan*, Ke Deng* and Nan Li* (2023). Robust estimation of dementia prevalence from a two-phase survey with missing responses via propensity score stratification. *BMC Medical Research Methodology* 23(1): 130.
- ◆ Feifei Huang, Lei Zhang, Moqin Zhou, Jianwen Li, Qing Liu, Bo Wang, Ke Deng, Pingping Zhou and Yongning Wu (2023). Polycyclic aromatic hydrocarbons in the chinese diet: Contamination characteristics, indicator screening, and health risk assessment. *Food Additives & Contaminants Part A* 40(5): 625-640.
- ◆ Nils Chr. Stenseth, Rudolf Schlatte, Xiaoli Liu, Roger Pielke, Jr., Ruiyun Li, Bin Chen, Ottar N. Bjørnstad, Dimitri Kusnezov, George Fu Gao, Christophe Fraser, Jason D Whittington, Yuqi Bai, Ke Deng, Peng Gong, Dabo Guan, Yixiong Xiao, Bing Xu and Einar Broch Johnson (2023). How to avoid a local epidemic becoming a global pandemic? *Proceedings of the National Academy of Sciences of United States of America* 120(10): e2220080120.
- ◆ Wanchuang Zhu, Yingkai Jiang, Jun S. Liu and Ke Deng* (2023). Partition-mallows model and its inference for rank aggregation. *Journal of the American Statistical Association* 118(541): 343-359.
- ◆ Yang Yang and Ke Deng* (2023). Generalized theme dictionary models for association pattern discovery. *The Annals of Applied Statistics* 17 (1): 269-293.
- ◆ Shengen Shawn Hu, Lin Liu, Qi Li, Wenjing Ma, Michael J. Guertin, Clifford A. Meyer, Ke Deng, Tingting Zhang and Chongzhi Zang (2022). Accurate estimation of intrinsic biases for improved analysis of chromatin accessibility sequencing data using SELMA. *Nature Communications* 13(1): 5533.
- ◆ Yicaho Li and Ke Deng* (2022). Simultaneous topology and loss tomography via theme dictionary models. *IEEE Transactions on Signal Processing* 70: 4239-4251.
- ◆ Feiyu Jiang, Dong Li, Wai Keung Li and Ke Zhu (2023). Testing and modelling for the structural change in covariance matrix time series with multiplicative form. *Statistica Sinica* 33(2): 787-818.
- ◆ Donghang Luo, Ke Zhu, Huan Gong and Dong Li* (2023). Testing error distribution by kernelized Stein discrepancy in multivariate time series models. *Journal of Business & Economic Statistics* 41(1): 111-125.
- ◆ Dong Li, Mui Li and Lianbin Zeng (2022). Simulation and application of subsampling for threshold



autoregressive moving-average models. *Communications in Statistics: Simulation and Computation* 51(5): 2110-2121.

◆ Xuanling Yang and Dong Li* (2022). Estimation of the empirical risk-return relation: A generalized-risk-in-mean model. *Journal of Time Series Analysis* 43(6): 938-963.

◆ Nils Chr. Stenseth*, Yuxin Tao, Chutian Zhang, Barbara Bramanti, Ulf Büntgen, Xianbin Cong, Yujun Cui, Hu Zhou, Lorna Dawson, Sacha Mooney, Dong Li, Henry Fell, Samuel Cohn, Florent Sebbane, Philip Slavin, Wannian Liang, Howell Tong, Ruifu Yang* and Lei Xu* (2022). No evidence for permanent natural plague reservoirs in historical and modern Europe. *Proceedings of the National Academy of Sciences of the United States of America* 119(51): e2209816119.

◆ Jiacheng Miao, Hanmin Guo, Gefei Song, Zijie Zhao, Lin Hou* and Qiongshi Lu* (2023). Quantifying portable genetic effects and improving cross-ancestry genetic prediction with GWAS summary statistics. *Nature Communications* 14: 832.

◆ Nayang Shan, Yao Lu, Hao Guo, Dongyu Li, Jitong Jiang, Linlin Yan, Jiudong Gao, Yong Ren, Xingming Zhao and Lin Hou* (2022). CITEDb: a manually curated database of cell-cell interactions in human. *Bioinformatics* 38(22): 5144-5148.

◆ Shuang Song, Hongyi Sun, Jun S. Liu* and Lin Hou* (2022). Multi-cell-type openness-weighted association studies for trait-associated genomic segments prioritization. *Genes* 13: 1220.

◆ Bo Yu, Pan Li, Qiangfeng Cliff Zhang* and Lin Hou* (2022). Differential analysis of RNA structure probing experiments at nucleotide resolution: uncovering regulatory functions of RNA structure. *Nature Communications* 13: 4227.

◆ Yucong Lin, Keming Lu, Sheng Yu, Tianxi Cai and Marinka Zitnik* (2023). Multimodal learning on graphs for disease relation extraction. *Journal of Biomedical Informatics* 143:104415.

◆ Zhengyun Zhao, Yichen Tian, Zheng Yuan, Peng Zhao, Feng Xia* and Sheng Yu* (2023). A machine learning method for improving liver cancer staging. *Journal of Biomedical Informatics* 137:104266.

◆ Qiao Jin, Zheng Yuan, Guangzhi Xiong, Qianlan Yu, Huaiyuan Ying, Chuanqi Tan, Mosha Chen, Songfang Huang, Xiaozhong Liu and Sheng Yu* (2023). Biomedical question answering: A survey of approaches and challenges. *ACM Computing Surveys* 55(2):1-36.

◆ Huaiyuan Ying, Shengxuan Luo, Tiantian Dang, and Sheng Yu* (2022). Label refinement via contrastive learning for distantly-supervised named entity recognition. In *Findings of the Association for Computational Linguistics: NAACL 2022*: 2656-2666.

◆ Hongyi Yuan, Zheng Yuan and Sheng Yu* (2022). Generative biomedical entity linking via knowledge base-guided pre-training and synonyms-aware fine-tuning. In *Proceedings of the 2022 Conference of the*

North American Chapter of the Association for Computational Linguistics: Human Language Technologies: 4038-4048.

- ◆ Xin Lu, Tianle Liu, Hanzhong Liu* and Peng Ding (2023). Design-based theory for cluster rerandomization. *Biometrika* 110(2): 467-483.
- ◆ Hanzhong Liu, Fuyi Tu and Wei Ma* (2023). Lasso-adjusted treatment effect estimation under covariate-adaptive randomization. *Biometrika* 110(2): 431-447.
- ◆ Xinhe Wang, Tingyu Wang and Hanzhong Liu* (2023). Rerandomization in stratified randomized experiments. *Journal of the American Statistical Association* 118(542): 1295-1304.
- ◆ Wei Ma, Fuyi Tu and Hanzhong Liu* (2022). Regression analysis for covariate-adaptive randomization: A robust and efficient inference perspective. *Statistics in Medicine* 41: 5645-5661.
- ◆ Ke Zhu and Hanzhong Liu* (2022). Confidence intervals for parameters in high-dimensional sparse vector autoregression. *Computational Statistics & Data Analysis* 168: 107383.
- ◆ Yicheng Li, Haobo Zhang and Qian Lin* (2023). On the Saturation effects of kernel ridge regression. *ICLR* 2023.
- ◆ Haobo Zhang, Yicheng Li, Weihao Lu and Qian Lin* (2023). On the optimality of misspecified kernel ridge regression. *ICML2023*.
- ◆ Subhra Sankar Dhar and Weichi Wu (2023). Comparing time varying regression quantiles under shift invariance. *Bernoulli* 29(2): 1527-1554.
- ◆ Tianying Wang*, Iuliana Ionita-Laza and Ying Wei (2022). Integrated quantile rank test (iqrat) for gene-level associations. *Annals of Applied Statistics* 16(3): 1423-444.
- ◆ Roulan Jiang, Xiang Zhan* and Tianying Wang* (2023). A flexible zero-inflated poisson-gamma model with application to microbiome read count data. *Journal of the American Statistical Association* 118(542): 792-804.
- ◆ Shuang Zhou, Debdeep Pati, Tianying Wang, Yun Yang and Raymond J. Carroll (2023). Gaussian processes with errors in variables: Theory and computation. *Journal of Machine Learning Research* 24: 1-53.
- ◆ Shiyang Ma and Tianying Wang* (2023). The optimal pre-post allocation for randomized clinical trials. *BMC Medical Research Methodology* 23:72.
- ◆ Jingyi Zhang, Cheng Meng, Jun Yu, Mengrui Zhang, Wenxuan Zhong and Ping Ma (2023). An optimal transport approach for selecting a representative subsample with application in efficient kernel density estimation. *Journal of Computational and Graphical Statistics* 32(1): 329-339.
- ◆ Lili Su, Jiaming Xu and Pengkun Yang (2022). Global convergence of federated learning for mixed



regression. *Advances in Neural Information Processing Systems* 35: 29889-29902.

- ◆ Natalie Doss, Yihong Wu, Pengkun Yang and Harrison H. Zhou (2023). Optimal estimation of high-dimensional Gaussian location mixtures. *The Annals of Statistics* 51(1): 62-95.
- ◆ Haoyue Tang, Yuchao Chen, Jintao Wang, Pengkun Yang and Leandros Tassioulas (2023). Age optimal sampling under unknown delay statistics. *IEEE Transactions on Information Theory* 69(2): 1295-1314.

In Press

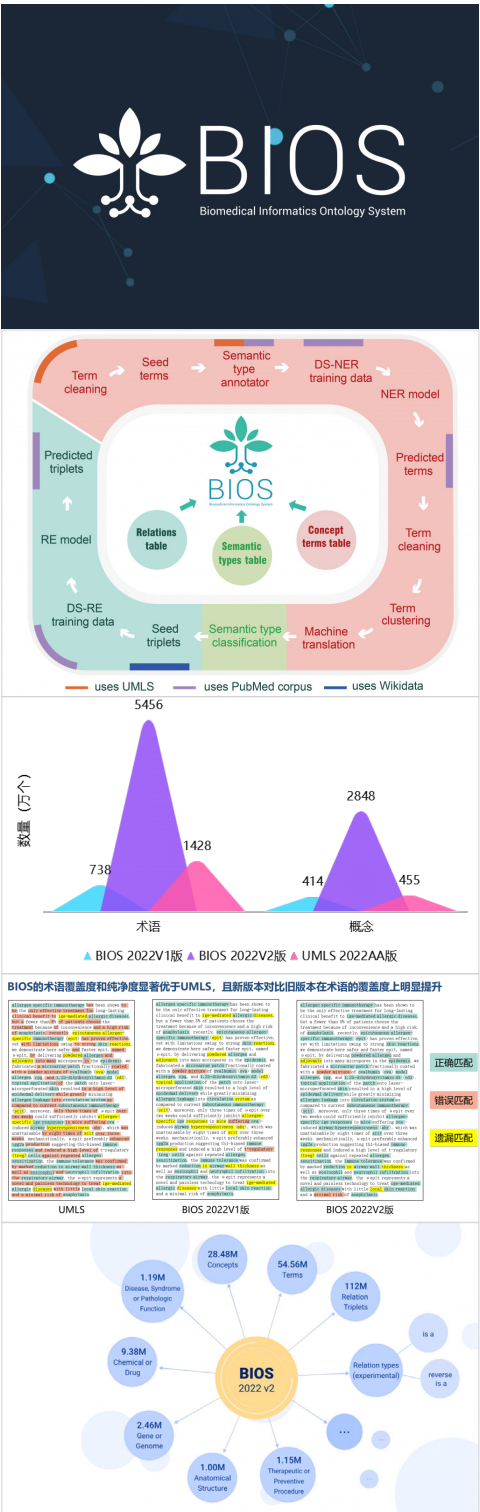
- ◆ Zening Song, Lijian Yang* and Yuanyuan Zhang (2023+). Hypotheses testing of functional principal components. *Statistica Sinica*.
- ◆ Chen Zhong and Lijian Yang* (2023). Statistical inference for functional time series: autocovariance function. *Statistica Sinica* 33(4).
- ◆ Xinyu Zhang, Dong Li* and Howell Tong (2023+). On the least squares estimation of multiple-threshold-variable autoregressive models. *Journal of Business & Economic Statistics*.
- ◆ Xuanling Yang, Zhoufan Zhu, Dong Li and Ke Zhu (2023+). Asset pricing via the conditional quantile variational autoencoder. *Journal of Business & Economic Statistics*.
- ◆ Yuxin Tao, Dong Li and Xiaoyue Niu (2023+). Grouped network Poisson autoregressive model. *Statistica Sinica*.
- ◆ Dong Li, Yuxin Tao, Yaxing Yang and Rongmao Zhang (2023). Maximum likelihood estimation for α -stable double autoregressive models. *Journal of Econometrics* 236(1): 105471.
- ◆ Yicheng Li, Haobo Zhang and Qian Lin* (2023). Kernel interpolation generalizes poorly. *Biometrika*.
- ◆ Yun Ma, Yihong Wu and Pengkun Yang (2023+). On the best approximation by finite Gaussian mixtures. In 2023 IEEE International Symposium on Information Theory (ISIT).
- ◆ Lili Su, Jiaming Xu and Pengkun Yang (2023+). A non-parametric view of FedAvg and FedProx: Beyond stationary points. *Journal of Machine Learning Research*.

教材出版：

《统计学基础》，高等教育出版社(2022)，郭建华、姚方、邹长亮、邓柯编著(同等贡献)



BIOS喜迎重大更新，成为全球最大开放生物医学知识图谱



清华大学统计学研究中心俞声课题组和粤港澳大湾区数字经济研究院 (IDEA) AI平台技术研究中心联合开发的大型开放生物医学知识图谱——“生物医学信息学本体系统” BIOS (Biomedical Informatics Ontology System) 迎来重大更新，跃升成为世界最大的开放生物医学知识图谱 (<https://bios.idea.edu.cn>)。

生物医学知识图谱是一种由生物医学概念、术语、关系以及 ID 系统等要素构成的特殊数据库，是生物医学信息学的重要基础设施。一直以来，由美国开发的“一体化医学语言系统” UMLS (Unified Medical Language System) 是生物医学知识图谱的标杆，以 455 万概念、2095 万关系的巨大规模和开放属性，为英文领域生物医药大数据分析、自然语言处理、人工智能开发和数据交换做出了卓越贡献。中文领域由于缺乏可开放获取的大型生物医学知识图谱，导致国内的医学大数据分析缺乏平台基础，科研与技术发展受到严重制约。同时，基于多数据库整合和专家整理的 UMLS 也日渐老化，其数据质量与发展速度已无法满足大数据与人工智能时代的需要。

基于算法驱动、开源开放理念，以 CC BY-NC-ND 4.0 协议发布的 BIOS 中英文知识图谱自 2021 年 11 月发布以来受到了广泛关注。BIOS 是首个完全由机器学习算法生成的大型开放生物医学知识图谱，其术语发现、语义分析、概念生成、关系发现、跨语言对齐完全由模型自动实现。在本次更新中，研发团队根据真实数据效果，不断强化算法技术，终于取得了振奋人心的突破性进展，使新版 BIOS (2022V2 版) 一举达到了 2848 万概念、5456 万术语 (3348 万英文、2108 万中文) 的巨大体量，术语质量也得到进一步提升。

而为了满足如此体量概念的关系挖掘，研究团队创新地提出了由“基于文本的关系提取”到“基于大模型自有知识的关系生成”的模式转变，获得了 1.12 亿个关系三元组。目前 BIOS 只包含知识图谱的骨架“上下位关系”的预测，更丰富的关系将在下半年的更新中发布。

对比开发已有 35 年的 UMLS，BIOS 在短短一年半的时间里，使体量达到了 UMLS 的数倍，不仅扭转了中文领域缺乏大型开放生物医学知识图谱的困难局面，更充分证明了人工智能的巨大潜力。同时，通过 BIOS 系统的研发，统计学研究中心培养了一批具有数据科学方向理论创新与实战能力的优秀本科生与博士生。未来，统

计学研究中心将与 IDEA 研究院以及更多国内外顶尖医院、科研机构合作，不断扩大和完善 BIOS 的内容、质量以及相关系统建设，带动并引领生物医学大数据与人工智能行业的发展。



统计中心18级博士研究生余博在 *Nature Communications* 发表论文开发 RNA结构差异分析方法



ARTICLE

<https://doi.org/10.1038/s41467-022-31875-3>

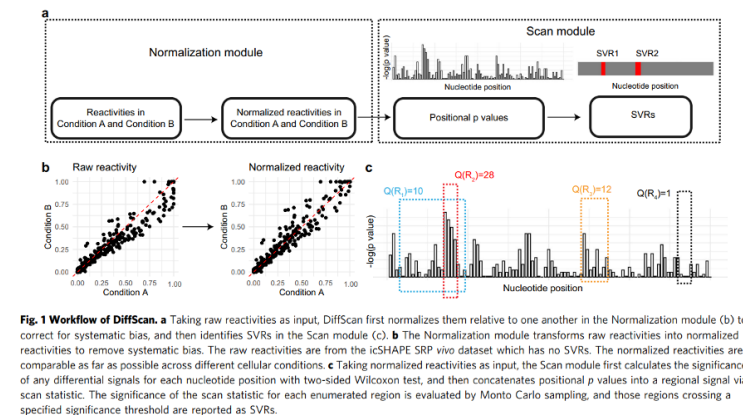
OPEN

Check for updates

Differential analysis of RNA structure probing experiments at nucleotide resolution: uncovering regulatory functions of RNA structure

Bo Yu¹, Pan Li^{2,3}, Qiangfeng Cliff Zhang^{2,3} & Lin Hou^{1,2}

“Differential analysis of RNA structure probing experiments at nucleotide resolution: uncovering regulatory functions of RNA structure” 的研究论文 (<https://www.nature.com/articles/s41467-022-31875-3>)。该研究提出了一个兼容多种 RNA 结构探测数据的标准化及结构差异分析的计算框架。该方法可以在单碱基分辨率水平评估 RNA 结构差异区域，为探索 RNA 结构在不同生物学过程中发挥的作用提供强有力的支持。同时，该研究还开发了相应的软件包 DiffScan (<https://github.com/yub18/DiffScan>) 供广大研究人员使用。清华大学统计学研究中心 2018 级博士研究生余博是该文章的第一作者，清华大学生命科学学院的博士毕业生李盼亦参与了该项研究。此外，该研究得到了国家自然科学基金和国家重点研发计划的支持。



大难度。在本研究中，研究团队提出针对高通量结构探测数据的标准化和差异分析计算框架 DiffScan，该方法首先对结构探测数据进行标准化以移除系统性偏差，再运用扫描统计量在 RNA 结构组中识别结构差异区域。

DiffScan 的独特贡献在于：(1) 其标准化模块能有效地移除实验间存在的系统性偏差，并充分保留结构差异信号；(2) 其扫描统计量能够在单碱基分辨率水平自适应地识别结构差异区域的位置与长度；(3) 该计算框架可以兼容包括 icSHAPE、DMS-seq 在内的多种高通量 RNA 结构探测技术产生的实验数据。

研究团队将 DiffScan 应用于不同亚细胞的 RNA 结构组数据中，通过结合相关的基序富集分析阐明了 RNA 结构变化与 mRNA 丰度之间的潜在联系可能是由包括 serine/arginine rich splicing factors 在内的特定 RNA 结合蛋白所介导的。因此，该研究也证明了 DiffScan 是破译 RNA 结构组数据的有效手段。

清华大学统计学研究中心侯琳副教授课题组与生命科学学院张强锋副教授课题组合作在 *Nature Communications* 期刊上在线发表了题为“Differential analysis of RNA structure probing experiments at nucleotide resolution: uncovering regulatory functions of RNA

RNA 通过形成特定的结构发挥相应的功能，并且 RNA 结构在不同的细胞环境中往往存在差异。因此，通过对不同条件下的 RNA 结构组数据进行准确的结构差异区域分析，对于揭示 RNA 在不同生物学过程中的功能及作用机制至关重要。考虑到高通量的 RNA 结构探索技术产生的实验数据往往存在高维度、高噪音等特点，RNA 结构的差异分析存在较



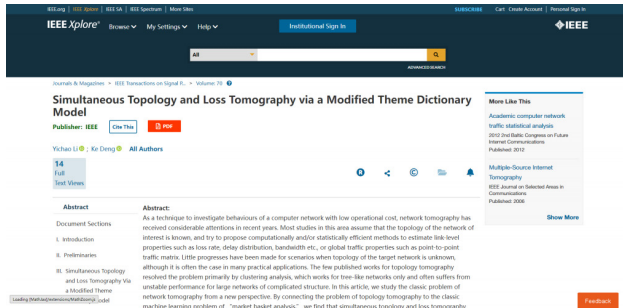
清华大学统计学研究中心刘汉中副教授课题组在国际知名统计学期刊 *Biometrika* 发表题为 “Design-based theory for cluster rerandomization” 的研究论文。清华大学统计学研究中心 2020 级博士研究生卢鑫是该文的第一作者，哈佛大学博士研究生刘天乐是第二作者，刘汉中副教授与加州大学伯克利分校的丁鹏副教授共同指导了相关研究与论文撰写。

群组随机化试验被广泛地应用在社会学、公共卫生等领域。它可以避免群组内部试验个体之间的交互，是在个体水平的处理分配不可实施的情形下一种替代的试验设计方法。群组随机化试验在群组的水平分配处理，同一个群组内部的个体会分配到同一个处理。群组重随机化是在群组随机化试验中运用重随机化的试验设计方法，通过丢弃掉不平衡的分配方式来达到平衡协变量的目的。重随机化最常用的平衡性准则有马氏距离准则，加权欧氏距离准则和多层马氏距离准则。马氏距离准则把各个协变量看成是同等重要的。当存在协变量重要程度的先验信息时，加权欧氏距离准则和多层马氏距离准则是更常用的方法。

该论文首先研究了群组重随机化下平均因果效应估计的渐近理论，证明了重随机化可以提高平均因果效应的估计精度。其次该论文比较了不同群组重随机化平衡性准则的效率，即在相同接受概率下对估计量渐近方差的减小量，证明了当协变量进行正交化之后，最优的加权欧式距离准则优于多层马氏距离准则。论文最后讨论了群组重随机化下的协变量回归调整方法，证明了在重随机化下，Su & Ding (2021)提出的模型辅助的点估计和区间估计仍然适用。这一结果对于因果推断的实践具有重要指导意义。



邓柯课题组在 *IEEE Transactions on Signal Processing* 发文提出用主题辞典模型分析网络行为



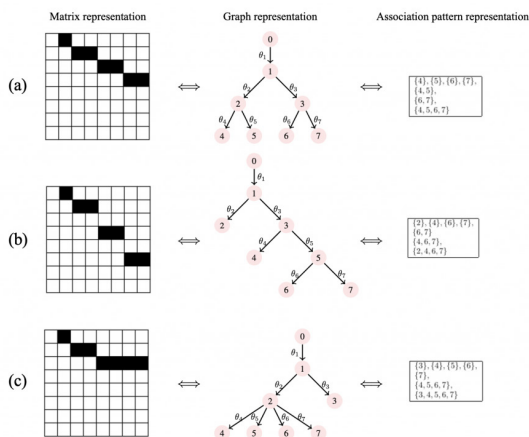
清华大学统计学研究中心邓柯副教授课题组在信号处理领域的顶级期刊 *IEEE Transactions on Signal Processing* 发表了题为“Simultaneous Topology and Loss Tomography via a Modified Theme Dictionary Model”的论文。通过对网络透视过程建立统计学模型，实现对目标网络的拓扑结构和丢包率的同时估计，并取得更准确、稳健的估计结果。课题组博士后研究员李艺超为第一作者，邓柯副教授为通讯作者。



李艺超 博士
第一作者



邓柯 副教授
通讯作者



通讯网络结构的三种等价表达形式：Matrix representation, Graph representation, Association pattern representation

进行估计的具体过程；(3) 在计算上，本文利用序贯蒙特卡洛(sequential Monte Carlo) 技术对 EM 算法中的 E 步进行了优化。

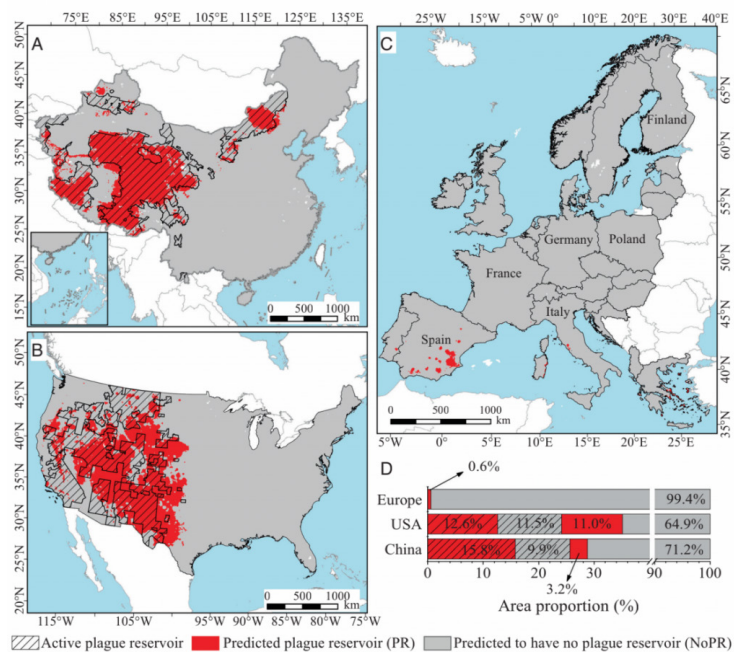
相对于传统的聚类方法，本文所提出的方法实现了对网络拓扑结构和各个节点上的丢包率进行同时估计；此外，在一些复杂网络中，本文所提出的方法展现出更准确、更稳健的估计结果。

该研究工作获得国家自然科学基金(Grant 11931001)、国家留学基金委以及清华大学国强研究院的支持。

现代社会中，我们的生活与网络息息相关。一些网络特征，例如网络拓扑结构、网络丢包率对我们研究网络行为非常重要。然而，出于安全、商业等原因的考虑，我们无法直接测量这些网络特征。为了克服这些困难，Vardi, Y. (1996)提出了网络透视技术(network tomography)。网络透视技术是一种兼具网络测试与网络预警的技术，旨在通过发送特定的数据包(探测包)到特定网络，利用端对端的方法将所得到的信息进行统计、综合分析从而进一步推断网络内部特征。

基于此，研究团队提出用统计学方法对网络透视过程建模，该项工作的主要贡献是：(1) 本文将网络丢包行为与统计学中的经典问题“购物篮分析”建立了对应关系：网络中的丢包行为可以和购物篮分析中的模式识别形成一一对应；(2) 进一步根据上述对应关系，本文利用主题辞典模型对网络丢包行为进行建模，并给出了利用 EM 算法对模型参数进行

李东课题组在 *PNAS* 合作发文证实欧洲不存在长期自然鼠疫源地的假设



鼠疫源地的空间预测分布

无论是在历史上还是在近现代，都没有证据支持欧洲自然鼠疫源地持续存在。有两个因素不利于欧洲鼠疫源地长期存在：(1) 土壤成分及生化性质；(2) 较低啮齿动物多样性。这些发现对于人类历史上的鼠疫研究具有广泛的意义，为解决围绕鼠疫长达一个世纪的谜团提供了新的工具。

清华大学统计学研究中心李东副教授课题组与清华大学杰出访问教授汤家豪院士、挪威奥斯陆大学 Nils Chr. Stenseth 院士、清华万科公共卫生与健康学院许磊教授研究团队等合作，在美国科学院院刊 *Proceedings of the National Academy of Sciences of the United States of America* (简称 PNAS) 发表了题为“*No evidence for persistent natural plague reservoirs in historical and modern Europe*”的论文。论文借助统计模型，证实了欧洲在历史和当代都不存在长期自然鼠疫源地的假设。2019级博士生陶宇心为该文的共同第一作者。

鼠疫(plague)是由鼠疫耶尔森菌(*Yersinia pestis*)借鼠蚤传播为主的烈性传染病。在过去两千年的三次大流行中，鼠疫夺去了数千万人的生命。关于鼠疫耶尔森氏菌是否持续存在于欧洲自然疫源地(natural plague reservoirs)中，还是从欧洲外部反复引入，始终是颇有争议的问题。

本文基于中国目前鼠疫源地的环境数据(包括土壤特征和气候)，利用统计中的MAVE方法进行统计建模，使用美国鼠疫源地相关数据验证模型的有效性和准确性，以评估欧洲的环境条件是否支持“天然鼠疫源地”的存在性。研究发现，无



统计中心 22级博士生江柔蓝在 JASA 发文提出零膨胀泊松-伽马模型分析微生物组数据



清华大学统计学研究中心王天颖助理教授课题组在国际知名统计学期刊 *Journal of the American Statistical Association* (简称“JASA”) 发表题为“A Flexible Zero-Inflated Poisson-Gamma Model with Application to Microbiome Sequence Count Data”的研究论文，提出了一种新的方法用于微生物组数据分析。清华大学统计学研究中心 2022级博士研究生江柔蓝为本文第一作者，中心王天颖助理教授和北京大学生物统计系占翔副教授为本文通讯作者。



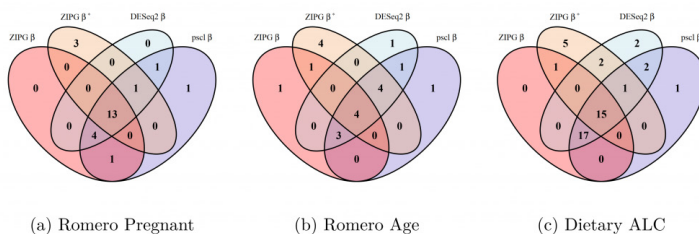
江柔蓝
第一作者

王天颖 助理教授
共同通讯作者

占翔 副教授
共同通讯作者

人体中生活着大量微生物，这些微生物群落受到人体环境的影响，也同时影响着人类的生命健康。近年来，飞速发展的高通量测序技术，如 16s rRNA 测序和宏基因组测序 (metagenomic shotgun sequencing)，为微生物群落的定量分析提供了便利。这些测序方法产生的数据以计数表的形式呈现，表示样本中各个微生物

物分类群 (taxa) 的丰度。微生物组数据往往具有零膨胀 (zero inflation) 和过度分散 (over-dispersion) 的特点，且受到人体环境的动态变化以及不同的测序技术的影响，观察到的计数数据往往与真实的微生物分类群丰度存在差异。为了研究某一微生物分类群与特定协变量的关联，现有的方法通常使用负二项分布或零膨胀负二项分布对计数数据建模，但是这些方法忽略了协变量对微生物群落稳定性的可能影响，尤其在纵向数据 (Longitudinal Data) 的分析中，已有证据表明疾病等因素会使微生物群落组成变得更不稳定，出现生态失调 (dysbiosis) 现象。为此，本文提出了零膨胀泊松-伽马模型 (Zero-Inflated Poisson-Gamma, ZIPG)，其中用伽马分布表示微生物组丰度相对无法观测到的真实均值的波动，能够更灵活地处理数据过度分散的特点。在纵向数据的分析中，通过将过度分散参数与时间无关 (time-independent) 的协变量关联起来，可以分析年龄、饮食习惯等长期状态对微生物组稳定性的影响。

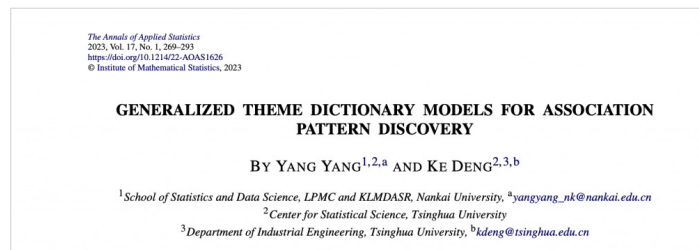


ZIPG与微生物数据主流分析方法的假设检验结果比较

在 ZIPG 的框架下，本文进一步提出了相应的参数估计和假设检验方法。通过 bootstrap 方法可以检验微生物分类群的

丰度和过度分散程度如何分别地受到协变量影响，并得到相应参数的置信区间。与其他方法相比，ZIPG 能够更好地控制第一类错误率，并且检测出其他方法无法发现的微生物稳定性差异。最后，本文将 ZIPG 应用在两组真实数据中：在阴道微生物组数据中分析了怀孕、年龄等因素对微生物丰度和过度分散程度的影响；在肠道微生物数据中分析了饮酒等因素对微生物丰度和过度分散程度的影响。在两例数据中，ZIPG 模型都能更好地拟合真实数据的分布，且找到了更多与感兴趣的协变量相关联的微生物分类群，为生物医学领域的进一步研究提供了线索。

邓柯课题组在 AOAS 发表论文提出用广义主题词典模型挖掘关联规则



清华大学统计学研究中心邓柯副教授课题组在应用统计知名期刊 *The Annals of Applied Statistics* (AOAS) 发表题为“Generalized theme dictionary models for association pattern discovery”的研究论文，提出了挖掘关联性规则的新方法。南开大学统计与数据科学学院杨洋特聘副研究员（课题组 2015 级博士生）为论文的第一作者，邓柯副教授为论文的通讯作者。

关联规则挖掘是数据挖掘研究的重要方向之一，广泛应用在生物信息学、医学和社会科学等领域。关联规则的一个典型应用是购物篮分析 (Market Basket Analysis, MBA)，即通过分析顾客购物篮中购买的商品，挖掘商品之间的关联性，实现智能推荐。在关联规则挖掘中，每个观测样本称为一个记录 (transaction)，每个记录是一组元素 (item) 的集合。以 MBA 问题为例，每件商品是一个

元素，每个购物篮构成的商品集合是一个记录。传统的关联规则挖掘方法不区分记录中元素的类别，并假设顾客购物篮中的商品数量无缺失，但随着应用场景的多元化，实际问题中的数据结构更为复杂，通常涉及两种类型的数据：在第一种类型的数据中，元素包含类别信息，传统方法可能挖掘出大量不具有较高应用价值的同类别关联规则；在第二种类型的数据中，元素的数量信息缺失，传统方法难以从缺失数据中正确挖掘关联规则。

该论文改进了 Deng, Geng and Liu (2014) 的方法，主要贡献如下：(1) 提出跨类别的主题词典模型，在产生初始词典的过程中限制主题的类别，对同类别的关联规则进行剪枝，能更准确、高效地挖掘跨类别的关联规则；(2) 提出联合主题词典模型，通过概率生成模型还原元素的数量信息，构造了一个高效的元素数量还原机制，以减少新模型增加的计算量；(3) 从理论上证明了新模型具有可识别性。与传统的关联规则挖掘方法相比，本文提出的方法在模拟实验和实际应用中都具有更高的计算效率、更高的准确性和更好的稳健性，并且该方法挖掘的关联规则在特征提取、命名体识别、聚类等下游任务中具有潜在的应用价值。

该研究工作获得国家自然科学基金 (Grant 11931001)、北京市自然科学基金 (Grant Z190021)、国家留学基金委、清华大学国强研究院以及教育部高等院校科技创新计划项目的支持。



- 专利及软著 -

【发明专利】



- 获批时间：2022.07.29
- 专利名称：医学关系提取方法及装置
- 发 明 人：俞声、林毓聪



- 获批时间：2022.10.14
- 专利名称：术语标准化方法及装置
- 发 明 人：俞声、余丽珊



- 获批时间：2022.11.15
- 专利名称：用于人工智能文本分析的双词典模型的初始化方法及设备
- 发 明 人：邓柯、徐嘉泽



- 获批时间：2022.11.29
- 专利名称：生理状态预测方法、计算机装置及存储介质
- 发 明 人：杨立坚、张园园、黄昆



- 获批时间：2022.12.02
- 专利名称：基于剖面隐式马氏模型的疾病名称标准化快速计算方法
- 发 明 人：邓柯、李祺、刘军



- 获批时间：2022.12.23
- 专利名称：电子病历标准化切分方法
- 发 明 人：俞声、苑洪意



- 获批时间：2023.01.24
- 专利名称：基于剖面隐式马尔科夫模型的疾病名称标准化方法和系统
- 发 明 人：邓柯、李祺、刘军



- 获批时间：2023.06.20
- 专利名称：基于卷积神经网络和连通图的图片归类方法及相关设备
- 发 明 人：邓柯、王海洋



- 获批时间：2023.06.27
- 专利名称：基于双词典模型的人工智能文本分析方法及相关设备
- 发 明 人：邓柯、徐嘉泽

【软件著作权】



- 盆腔器官脱垂的治疗方案智能推荐平台前端系统V1.0
- 获批时间：2022年12月16日
- 著作权人：清华大学、中国医学科学院北京协和医院
- 研发团队：邓柯、刘朝阳、朱兰、孙智晶



- RNA结构差异分析软件
- 获批时间：2023年1月29日
- 著作权人：清华大学
- 研发团队：余博、侯琳、张强锋

- 奖励荣誉 -



杨立坚 教授



2022年清华大学优秀博士学位指导教师



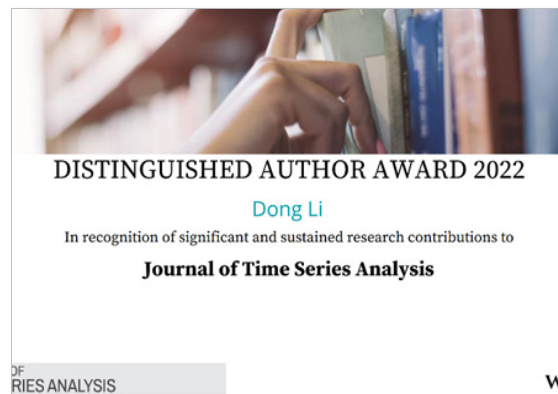
邓柯 副教授



邓柯副教授与清华大学自动化系生物信息学张学工教授团队联合报送的“基因差异表达及调控的模式识别”项目荣获 2022年度教育部“高等学校科研优秀成果奖”(科学技术类) 自然科学奖一等奖



李东 副教授



李东副教授荣获 2022年《时间序列分析》杂志 2022年度杰出作者称号



邓婉璐 副教授



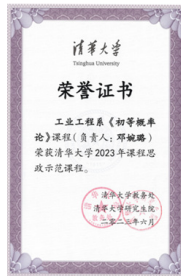
清华大学第五批标杆课
清华大学第九届“清韵
烛光·我最喜爱的教师”

表2 2022届毕业生心目中的好教师及课程							
序号	好教师名	课程号	课程名	问卷数	提名率	专业	
1.	邓婉璐	数据结构	00240074	116	29	24.88%	全校选课学生
2.	艾振华	高等微积分(I)	30420095	170	29	16.98%	全校选课学生
3.	李国林	电子电路与系统基础(1)	30230812	306	29	9.43%	全校选课学生
4.	王红	数字电子技术基础	20250103	246	24	30.57%	全校选课学生
5.	邱睿	辐射防护及保健物理	40320172	69	21	30.57%	全校选课学生
6.	李燕	中国近代史纲要	10610193	332	21	6.35%	全校选课学生
7.	卢旭光	测度与积分	30420334	58	20	34.45%	全校选课学生
8.	邓婉璐	初等概率论	40160713	117	20	17.08%	全校选课学生
9.	唐宏烈	概率论与数理统计	10420803	341	20	5.86%	全校选课学生
10.	李静杰	佛学概论文化	30807453	16	4	25%	艺术史论
11.	谭海霞	国际关系理论基础	30700362	14	5	35.71%	环境工程(全球环境国际班)
12.	杜月	西方社会学思想史下	40700183	20	5	25%	社会学
13.	王莹	语言心理学	40700642	37	11	29.73%	心理学

2022届毕业生心目中的好教师及课程



2021年度清华大学年度教学优秀奖
2022年度清华大学年度教学优秀奖



2023年课程思政示范课、课程
思政示范教师



清华大学精品课



周在莹 副教授



周在莹 副教授
清华大学 2022年度“青年教师教学优秀奖”



2021年度清华大学年度教学优秀奖
2022年度清华大学年度教学优秀奖



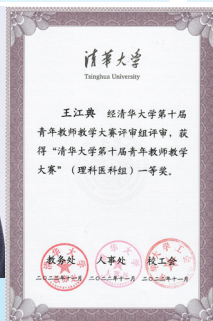
2022年课程思政示范课、课程
思政示范教师



清华大学精品课



王江典 讲师



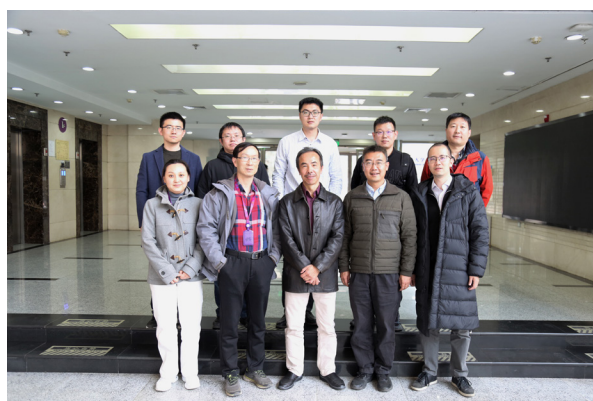
王江典 讲师
清华大学第十届青年教师教学大赛(理科、
医科组)一等奖

学术活动

- 主办学术活动 -

清华大学统计学研究中心-香港中文大学(深圳)数据科学学院交流研讨会

2023年3月20日,香港中文大学(深圳)数据科学学院姚建峰教授、于天维教授、刘瑾副教授、宋方达助理教授到访清华大学统计学研究中心,并出席“清华大学统计学研究中心-香港中文大学(深圳)数据科学学院交流研讨会暨统计学前沿研讨会”,会议由中心李东副教授发起并主持。双方研究团队在会上交流各自研究方向、分享最新研究成果,并针对未来可能合作的方向展开讨论。



机器学习基础理论研讨会

2023年3月25日,由清华大学统计学研究中心、清华大学交叉信息研究院与微软亚洲研究院联合主办的机器学习基础理论研讨会(Workshop on Machine Learning Theory and Foundations)在微软大厦顺利举办。本次研讨会邀请15名全球范围内机器学习领域的优秀学者,围绕最新的研究成果进行主题报告与分享,并针对当前机器学习领域的前沿问题展开深入交流与探讨。



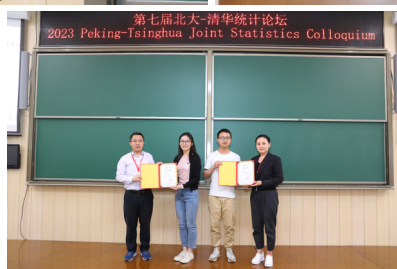
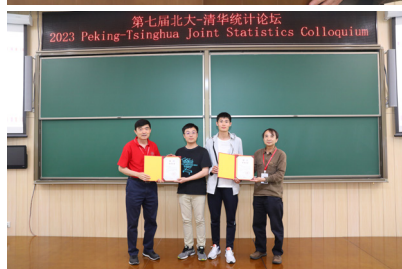


第七届北大-清华统计论坛

2023年5月25日，第七届北大-清华统计论坛在北京大学成功举办。论坛由北京大学统计科学中心和清华大学统计学研究中心联合发起，并作为兄弟院校间持续的品牌活动，每年定期举行。本届论坛线上线下累计有近二百人参会，除清北两校师生外，还有很多其他高校和业界的学者关注及参与。

清华大学苏良军教授和北京大学的张志华教授分别代表两校作大会特邀报告。论坛产生本年度“优秀毕业生”获得者：北京大学数学科学学院18级博士

研究生李翔、清华大学统计学研究中心18级博士研究生余博；“优秀海报奖”获得者：北京大学数学科学学院20级博士研究生帅康、北京大学光华管理学院20级博士研究生凌波、清华大学统计学研究中心20级博士研究生卢伟灏和清华大学统计学研究中心21级博士研究生马运。



第二十四届京津冀青年概率统计学术研讨会



2023年6月10日，由中国数学会概率统计学会主办、清华大学数学科学系和清华大学统计学研究中心共同承办的“第二十四届京津冀地区青年概率统计学术研讨会”在清华大学成功举办。来自北京大学、清华大学、中国科学院、中国科学院大学、中国人民大学、南开大学、北京师范大学、北京理工大学、首都师范大学、天津工业大学、河北科技大学等京津冀地区38所高校与科研机构的专家学者、青年教师、研究生等二百余人参加了会议。

- 科研项目 -

项目来源	项目类型	项目期限	项目金额	负责人
国家自然科学基金	面上项目	2022 年—2026 年	51 万 (元)	杨立坚
国家自然科学基金	数学天元基金项目	2021 年—2022 年	20 万 (元)	杨立坚
索元生物医药公司	合作项目	2021 年—2022 年	50 万 (元)	杨立坚
北京市机关事务管理局	国家机关委托项目	2023 年—2023 年	10 万 (元)	邓 柯
海关总署国际检验检疫标准与技术法规研究中心	国家机关委托项目	2022 年—2022 年	160 万 (元)	邓 柯
清华大学国强人工智能研究院	高校资助	2021 年—2023 年	100 万 (元)	邓 柯
北京市食品安全风险评估中心	国家机关委托	2021 年—2022 年	21 万 (元)	邓 柯
海关总署国际检验检疫标准与技术法规研究中心	国家机关委托项目	2021 年—2022 年	33 万 (元)	邓 柯
广西产品质量检验研究院	国家机关委托项目	2021 年—2022 年	10 万 (元)	邓 柯
国家自然科学基金	重点项目课题	2020 年—2024 年	50 万 (元)	邓 柯
科技部	国家重点研发计划子课题	2020 年—2023 年	54 万 (元)	邓 柯
北京市自然科学基金	重点项目课题	2019 年—2023 年	120 万 (元)	邓 柯
国家社会科学基金	重点项目课题	2018 年—2022 年	10 万 (元)	邓 柯
国家自然科学基金	面上项目	2020 年—2023 年	48 万 (元)	李 东
国家自然科学基金	面上项目	2021 年—2024 年	51 万 (元)	侯 琳
科技部	国家重点研发计划子课题	2020 年—2025 年	50 万 (元)	侯 琳
国家高层次人才特殊支持计划	青年拔尖人才	2023 年—2026 年	200 万 (元)	俞 声
国家自然科学基金	面上项目	2022 年—2025 年	50 万 (元)	俞 声



项目来源	项目类型	项目期限	项目金额	负责人
粤港澳大湾区数字经济研究院（福田）	合作研究	2021 年—2024 年	223.88 万（元）	俞 声
北京市自然科学基金	重点项目课题	2019 年—2023 年	100 万（元）	俞 声
国家自然科学基金	面上项目	2021 年—2024 年	52 万（元）	刘汉中
清华大学国强研究院	高校资助	2021 年—2023 年	100 万（元）	刘汉中（参与）
国家高层次人才特殊支持计划	青年拔尖人才	2022 年—2024 年	120 万（元）	林 乾
字节跳动	企业委托项目	2022 年—2024 年	45 万（元）	林 乾
北京市智源人工智能研究院	青年科学家专项	2022 年—2022 年	50 万（元）	林 乾
国家自然科学基金	面上项目	2020 年—2024 年	52 万（元）	林 乾
北京市自然科学基金	重点项目课题	2019 年—2023 年	65 万（元）	林 乾
国家自然科学基金	面上项目	2023 年—2026 年	46 万（元）	吴未迟
国家自然科学基金	青年项目	2019 年—2022 年	28.9 万（元）	吴未迟
国家自然科学基金	青年项目	2021 年—2024 年	30 万（元）	王天颖
科技部	国家重点研发计划青年科学家项目	2021 年—2026 年	12.8 万（元）	王天颖（参与）
北京医院	医工结合专项	2022 年—2024 年	12 万（元）	张静怡
科技部	国家重点研发计划青年科学家项目	2021 年—2026 年	28.5 万（元）	张静怡（参与）
北京市自然科学基金	面上项目	2023 年—2025 年	20 万（元）	张静怡（参与）
国家自然科学基金	优秀青年科学基金项目（海外）	2022 年—2025 年	100 万（元）	杨朋昆
国家自然科学基金	青年项目	2022 年—2024 年	30 万（元）	杨朋昆
清华大学自主科研计划	前沿交叉专项	2022 年—2023 年	80 万（元）	杨朋昆

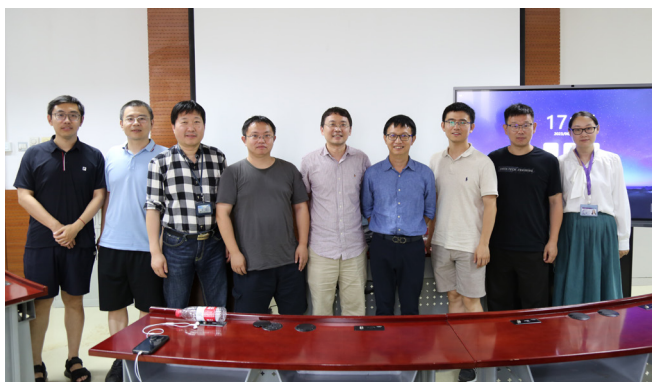


- 统计学与数据科学论坛 -

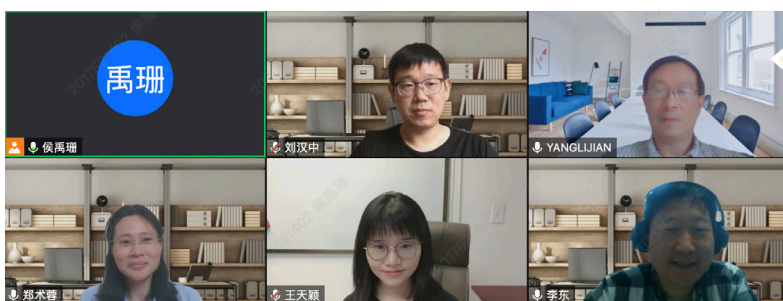
时间	主讲人	工作单位及职位	报告题目
2022.08.30	史成春	伦敦政经学院助理教授	Statistical Inference in Reinforcement Learning
2022.09.19	于 怡	华威大学 Reader	On Robustness and Local Differential Privacy
2022.09.26	谢 睿	中佛罗里达大学助理教授	Optimal Sampling Designs for Online Estimation of Streaming Multi-dimensional Time Series
2022.10.07	李 建	清华大学副教授	Optimization, Generalization and Implicit bias of Gradient Methods in Deep Learning
2022.10.14	顾雨琦	哥伦比亚大学助理教授	Bayesian Pyramids: Identifiable Multilayer Discrete Latent Structure Models for Discrete Data
2022.10.17	丁 剑	北京大学教授	Combinatorial Statistics: a Common Theme and a Few Examples
2022.10.24	Denis Belomestny	杜伊斯堡 - 埃森大学教授	Exploration in Reinforcement Learning via Randomisation
2022.10.31	凌渥丹	威尔康奈尔医学院	Statistical Analysis of Large-scale Microbiome-profiling Studies: Batch Effect and Robust Testing
2022.11.07	Grace Y. Yi	西安大略大学教授	How Myths about Noisy Data may Mislead Us
2022.11.14	时丕旭	杜克大学助理教授	Dimension Reduction of Longitudinal Microbiome Data by Tensor Functional SVD
2022.11.21	刘 歆	中国科学院研究员	Decentralized Optimization Over the Stiefel Manifold by an Approximate Augmented Lagrangian Function
2022.11.28	刘耀午	西南财经大学教授	A Power-robust Test for Global Hypotheses in Generalized Linear Models
2022.12.05	朱利平	中国人民大学教授	统计学与人工智能：融通共进
2022.12.12	郭 旭	北京师范大学副教授	Model-Free Statistical Inference on High-Dimensional Data
2022.12.19	乔兴昊	伦敦政经学院副教授	On the Modelling and Prediction of High-dimensional Functional Time Series
2023.02.13	Harrison Zhou	耶鲁大学教授	Leave-one-out Singular Subspace Perturbation Analysis for Spectral Clustering
2023.02.20	张 炘	佛罗里达州立大学副教授	Tensor Modeling in Categorical Data Analysis and Association Studies



时间	主讲人	工作单位及职位	报告题目
2023.02.27	於 州	华东师范大学教授	神经网络遇见随机森林
2023.03.06	焦雨頔	武汉大学副教授	深度学习的理论探究：逼近、估计、优化与表示
2023.03.13	张新雨	中科院数学与系统科学研究院研究员	Optimal Weighted Random Forests
2023.03.20	姚建峰	香港中文大学（深圳）教授	Weight Matrices Spectra in Deep Learning and Early-stopping
2023.03.27	张菁菲	埃默里大学副教授	Network Community Detection: New Algorithms and Goodness-of-fit Tests
2023.04.03	姚 方	北京大学教授	Theory of FPCA for Discretized Functional Data
2023.04.10	杨 灿	香港科技大学副教授	Strengthen Causal Inference Accounting by Leveraging Genetic Data
2023.04.17	周 舟	多伦多大学教授	Time-frequency Analysis, a Non-stationary Time Series Perspective
2023.04.20	林华珍	西南财经大学教授	Functional Data Analysis with Covariate-dependent Mean and Covariance Structures
2023.04.24	崔逸凡	浙江大学百人计划研究员	Some Recent Progress in Proximal Causal Learning
2023.05.08	程 豪	中国科协创新战略研究院副研究员	Fractional Imputation Algorithms in Quantile Regression
2023.05.22	骆 威	浙江大学百人计划研究员	A Unified Generalization of Inverse Regression via Adaptive Column Selection
2023.05.22	马诗洋	上海交通大学助理研究员	Knockoff-Based Statistics for the Identification of Putative Causal Genes in Genetic Studies
2023.05.29	郑术蓉	东北师范大学教授	高维尺度不变的判别分析
2023.06.05	夏应存	新加坡国立大学教授	Ensemble Projection Pursuit for General Nonparametric Regression
2023.06.19	张 凯	北卡罗来纳大学教堂山分校副教授	BET and BELIEF



2023年6月19日，新加坡国立大学夏应存教授访问中心



2023年6月5日，东北师范大学
郑术蓉教授通过线上平台与中心
教员交流



2023年4月20日，西南财经大学林华珍教授访问中心



2023年4月17日，多伦多大学周舟教授访问中心



2023年4月3日，北京大学姚方教授访问中心



2023年3月20日，香港中文大学（深圳）姚建峰教授访问中心



2023年3月13日，中国科学院数学与系统科学研究院张新雨研究员访问中心



2023年2月13日，耶鲁大学Harrison Zhou教授访问中心



2022年12月5日，中国人民大学朱利平教授通过线上平台与中心教员交流



2022年11月21日，中国科学院刘歆研究员访问中心



2022年10月24日，杜伊斯堡-埃森大学 Denis Belomestny教授通过线上平台与中心教员交流



2022年11月7日，西安大略大学Grace Y. Yi教授通过线上平台与中心教员交流



2022年10月17日，北京大学丁剑教授访问中心



- 参加学术活动 -

杨立坚

- 2023年04月 中国·北京 中国人民大学统计学院（邀请报告）
- 2023年04月 中国·线上 安徽大学大数据与统计学院（邀请报告）
- 2022年12月 中国·线上 浙江工商大学统计与数学学院（邀请报告）
- 2022年12月 中国·线上 武汉大学资源与环境科学学院第二届空间全生命周期健康国际研讨会（邀请报告）
- 2022年12月 中国·线上 西南交通大学数学学院（邀请报告）
- 2022年08月 中国·线上 四川大学数学学院国家自然科学基金（数学天元基金）西部高校数学教师暑期学校（邀请报告）

邓 柯

- 2023年08月 中国·湖州 湖州师范大学理学院
- 2023年07月 中国·北京 首届全国统计与数据科学联合会议
- 2023年07月 中国·开封 河南大学数学与统计学院
- 2023年07月 中国·成都 The 2023 ICSA China Conference
- 2023年04月 中国·青岛 第十二届全国概率统计学术会议（中会邀请报告）
- 2022年11月 中国·北京 第十五届中国R会议（大会邀请报告）
- 2022年11月 中国·上海 第五届中国国际进口博览会“非关税贸易措施高质量发展论坛”
- 2022年11月 中国·上海 华东师范大学统计学院
- 2022年09月 中国·西安 国家天元数学西北中心
- 2022年09月 中国·北京 2022中国国际服务贸易交易会“技术贸易措施助力经济高质量发展和国门生物安全建设高峰论坛”
- 2022年08月 中国·杭州 浙江大学数据科学研究中心
- 2022年08月 中国·成都 西南财经大学统计学中心

李 东

- 2023年06月 中国·深圳 中国现场统计研究会多元分析应用专业委员会学术年会
- 2023年04月 中国·上海 上海财经大学统计与管理学院（邀请报告）
- 2023年04月 中国·上海 复旦大学管理学院（邀请报告）
- 2023年04月 中国·青岛 第12届全国概率统计会议
- 2023年03月 中国·北京 北京大学光华管理学院（邀请报告）
- 2022年12月 中国·厦门 厦门大学统计会议（邀请报告）
- 2022年11月 中国·西宁 青海师范大学数学与统计学院（邀请报告）
- 2022年07月 中国·长春 吉林大学数学学院（邀请报告）
- 2022年07月 中国·西安 西安交通大学数学与统计学院（邀请报告）

侯 琳

- 2023年07月 中国·济南 第9届国际统计遗传与基因组学高峰论坛（邀请报告）
- 2023年05月 中国·香港 International Conference on Applied Mathematics 2023（邀请报告）
- 2023年05月 中国·徐州 第9届全国计算生物学与生物信息学学术会议（邀请报告）
- 2023年02月 中国·广州 生物信息学与智能信息处理2022年学术年会（邀请报告）
- 2022年12月 中国·线上 第七届中国计算机学会生物信息学会议（邀请报告）

俞 声

- 2022年08月 中国·线上 中国现场统计研究会数据科学与人工智能分会年会（大会报告）
- 2022年07月 中国·线上 北京大学人民医院（邀请报告）

刘汉中

- 2022年09月 中国·北京 The 2022 Pacific Causal Inference Conference（分会场报告）

林 乾

- 2023年07月 中国·成都 ICSA2023（邀请报告）
- 2023年07月 中国·北京 数据科学与统计联合会议（邀请报告）
- 2023年06月 中国·香港 ICAM2023（邀请报告）
- 2023年03月 中国·上海 全国青年统计学家含计量保险论坛（邀请报告）

吴未迟

- 2023年06月 中国·香港 香港中文大学统计系
- 2023年06月 中国·香港 香港中文大学（深圳）数据科学学院（邀请报告）
- 2023年02月 中国·杭州 浙江大学数据科学研究中心
- 2022年09月 法国·雷恩 5TH Workshop on GOODNESS-OF-FIT, CHANGE-POINT and Related Problems

张静怡

- 2023年06月 中国·深圳 南方科技大学医学院（邀请报告）
- 2023年06月 中国·深圳 香港中文大学（深圳）数据科学学院（邀请报告）
- 2022年07月 中国·武汉 武汉大学数学与统计学院（邀请报告）

杨朋昆

- 2023年06月 中国·线上 ISIT2023（发言）
- 2023年05月 中国·上海 华东师范大学统计学院（邀请报告）
- 2023年04月 中国·青岛 第十二届全国概率统计会议（发言）
- 2023年03月 中国·成都 西南财经大学统计研究中心（邀请报告）
- 2023年03月 中国·北京 机器学习理论与基础模型研讨会（组织者）
- 2022年11月 美国·线上 NeurIPS 2022（发言）
- 2022年09月 中国·上海 上海交通大学数学科学学院（邀请报告）
- 2022年09月 美国·线上 Allerton 2022（邀请报告）
- 2022年08月 中国·北京 中国机器学习与科学应用大会（邀请报告）

邓婉璐

- 2023年06月 中国·线上 统计学课程思政虚拟教研室第二次会议（邀请报告）
- 2023年05月 中国·成都 2023工业工程专业基础课程虚拟教研室研讨会（邀请报告）
- 2023年05月 中国·北京 清华大学教师发展中心“教与学”教师分享会（邀请报告）



人才培养 - 本科生培养 -

统计学专业辅修学位

自 2016 年开始，清华大学统计学研究中心开设统计学辅修项目，为学有余力的清华学子开设优质统计学课程。课程设计参考国际一流统计学科本科生培养方案，理论和应用并重，力争培养兼具统计学理论基础和应用能力的跨学科人才。近年来，辅修项目受到同学们的热烈欢迎，开设课程供不应求，多次响应同学们呼吁增加课程容量，年度新增课程修读申请超二百人次。2021 年年初，清华大学工业工程系统统计学本科专业通过教育部审批备案，辅修统计学专业并满足培养方案培养要求的同学可授予统计学辅修专业学位。

课程编号	课程名称	学分	学期	先修要求
必修课程 16 学分				
40160713	初等概率论	3	秋	微积分、线性代数
30160263	统计推断	3	秋	微积分、线性代数
40160763	多元统计分析	3	春	统计推断、初等概率论
40160803	线性回归分析	3	春	统计推断、初等概率论
30160294	统计计算与软件	4	秋	统计推断、初等概率论



“数据思维与实践”课程证书项目

自 2020 年开始，清华大学统计学研究中心开设“数据思维与实践”课程证书项目。项目对修读学生的专业及年级要求放宽，以满足对统计学科有热情但无法达到辅修学位要求的同学们的修读需求。项目必修课 2 门，共 6 学分；选修课从课组中的 15 门课程自由选择修读，学分要求不少于 9 学分。完成项目要求并通过资格审查的学生可授予课程项目证书。

通识课建设：“统计学引论：数据分析的科学与艺术”

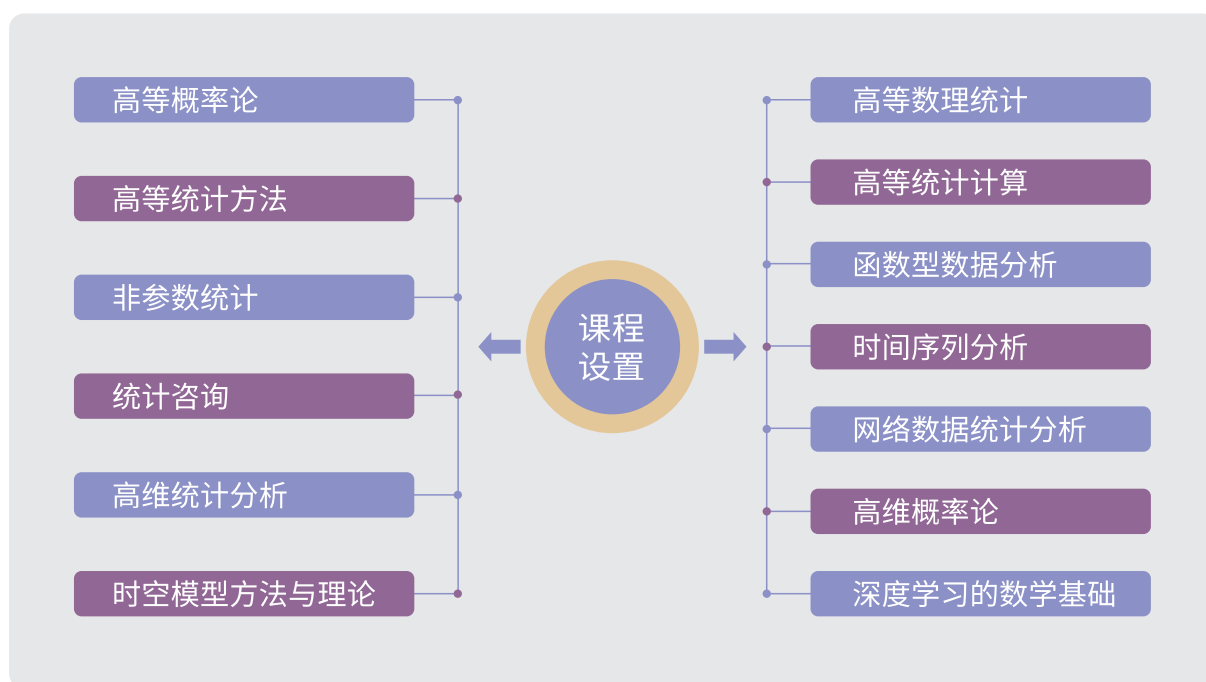
2018 年，清华大学统计学研究中心对标国际顶尖院校的统计学通识课程，筹备建设了清华大学首门统计学通识课，以应对大数据时代各专业领域对统计学的强烈需求。课程不但面向全校各专业本科生开放，更加入清华 - 北大互选课名单，预留部分名额，向兄弟院校的学生开放。课程由中心长聘副教授邓柯、李东两位老师讲授。内容包括统计学基本思想、基础理论和基本方法的理论，横跨工程、生物医学、经济金融、人文社科等多个专题，融合数据分析实践和统计编程入门的实践。



- 研究生培养 -

统计学博士

培养德智体全面发展，掌握扎实统计学基础理论和系统专业知识，具有独立从事统计学原创性研究和应用能力的统计学人才；使学生具有统计学素养，掌握学术规范，具备独立开展学术研究和进行学术交流的能力；指导学生应用统计学知识解决实际问题，在有关的统计学研究方向上做出重要理论价值或实际应用的创新性成果；毕业以后，适合在高等学校、科研机构、政府部门、企事业单位中从事统计学及其相关领域的教学、科研、管理等方面的研究和工作的。主要研究方向：数理统计、生物与医学统计、计量经济与金融统计、大数据统计、工业统计、统计计算等。





博士生论坛

2022年11月,由清华大学统计学研究中心发起并组织的“2022年清华大学统计学博士生论坛”成功举办。统计学博士生论坛是清华大学统计学研究中心的传统学术活动,旨在为青年统计学者提供学术交流的平台,以提高统计学者的专业知识及专业素养。按照清华大学统计学博士研究生培养方案,中心二年级及以上的博士生每年都要汇报自己的研究进展。

来自清华大学统计学研究中心的四十余名在读博士生参与了此次论坛,青年学者们根据各自的研究方向,分享最新研究成果以及在研究中遇到的问题,切磋技艺,相互交流,受益匪浅。



学生获奖情况

奖项名称	获奖者	具体信息	时间	导师
 第八届全国高校 研究生论坛十佳 论文	宋泽宁	Hypotheses Testing of Functional Principal Components	2022 年 11 月	杨立坚
 北京生物医学统计 与数据管理研究会 2023年度“百济 神州”青年优秀论 文特等奖	沈 翀	Model-Based Spatial Reconstruction of Large-Scale Biomolecules via Bayesian Inference of a Hierarchical Spatial Model	2023 年 6 月	邓柯
 2023 Hannan Graduate Student Travel Award (2023 年汉南研究 生旅行奖)	陶宇心	Statistical Inference for Stable Asymmetric GARCH Models	2023 年 4 月	李东
第六届全国统计学 博士研究生学术论 坛优秀论文二等奖	陶宇心	Grouped Network Poisson Autoregressive Model	2022 年 11 月	李东
北京应用统计学会 学术研讨会优秀论 文奖	陶宇心	Maximum Likelihood Estimation for Alpha-stable Double Autoregressive Models	2022 年 9 月	李东
 第二十四届京津 冀青年概率统计 学术研讨会“钟家 庆优秀论文奖”	宋 爽	A Robust Penalized-regression- based Method for Multivariable Mendelian Randomization Using GWAS Summary Statistics	2023 年 6 月	刘军 侯琳
 第七届北大 - 清华 统计学论坛优秀 毕业生	余 博	/	2023 年 6 月	侯琳



奖项名称	获奖者	具体信息	时间	导师
中国生物统计 2022 年学术年会优秀论文三等奖	余 博	Differential Analysis of RNA Structure Probing Experiments at Nucleotide Resolution: Uncovering Regulatory Functions of RNA Structure	2022 年 8 月	侯琳
 北京生物医学统计与数据管理研究会 2023 年度“百济神州”青年优秀论文二等奖	朱 珂	Design-based Theory for Lasso Adjustment in Randomized Block Experiments and Rerandomized Experiments	2023 年 6 月	刘汉中
第六届全国统计学博士研究生学术论坛优秀论文三等奖	朱 珂	Design-based Theory for Lasso Adjustment in Randomized Block Experiments with a General Blocking Scheme	2022 年 11 月	刘汉中
北京应用统计学学会学术研讨会优秀论文奖	朱 珂	Design-based Theory for Lasso Adjustment in Randomized Block Experiments with a General Blocking Scheme	2022 年 9 月	刘汉中
 北京应用统计学学会学术研讨会优秀报告奖	张皓博	Variance Reduction in Online A/B Testby Inference on Ex-post Residual	2022 年 9 月	林乾
 第七届北大 -清华统计学论坛优秀优秀海报奖	卢伟灏	Optimal Rate of Kernel Regression in Large Dimensions	2023 年 6 月	林乾
 第七届北大 -清华统计学论坛优秀优秀海报奖	马 云	Random Smoothing Regularization in Kernel Gradient Descent Learning	2023 年 6 月	杨朋昆

优秀毕业生专栏

2022年7月—2023年6月，清华大学统计学研究中心共8名同学顺利毕业：黄昆、宋泽宁、沈翀、杨萱铃、余博、罗声旋、罗赛迪、朱珂。



清华大学统计学研究中心博士生风采

黄 昆

导师：杨立坚 / 毕业去向：德州农工大学博士后



感谢杨立坚老师引领我进入科研世界的大门。科研是一个痛并快乐着的过程，其中会遇到各种各样的困难，需要我们有足够的勇气与毅力去面对。回想起来，这五年里我经历了遇到问题时的彷徨失措，思考问题时的辗转反侧，证明定理时的绞尽脑汁，直到最后解决问题后的悠然自得。这些磨炼也让我逐渐成长。回首往事，我会感谢这些年在统计中心度过的岁月。感谢统计中心的老师和同学们。也希望还没毕业的学弟学妹们好好珍惜在学校的生活，人生苦短，道路漫长，要积极进取，把握当下。最后，祝统计中心发展得越来越好！

宋泽宁

导师：杨立坚 / 毕业去向：南开大学讲师



转眼五年的博士生涯已经结束，在清华的学习经历让我受益良多。首先非常感谢统计中心为我们提供了雄厚的师资力量、丰富的学术资源和良好的科研氛围，感谢中心每一位老师的教导和所有同学的陪伴。在此也要特别感谢我的导师杨立坚教授，杨老师在学习和生活方面给予了我无微不至的关怀，他严谨求实的治学态度、扎实深厚的学识功底、精益求精的工作作风，都是我终身学习的榜样。最后，衷心祝愿统计中心越办越好！

沈 翀

导师：邓柯 / 毕业去向：某部委



感谢统计学研究中心为我们提供的高水平学术平台，在这里不仅有丰富的学术资源，还有自由的交流氛围。感谢中心各位老师的教导和帮助，特别感谢邓老师一直以来的授业解惑、循循善诱，让我在各方面都受益匪浅。难忘一路走来各位同学的支持和陪伴，衷心祝愿统计中心越办越好！



杨萱铃

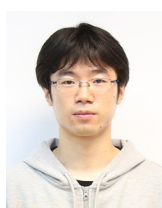
导师：李东 / 毕业去向：中金财富证券



在中心的五年时光，充实又温暖。读博很奇妙：求学时感觉不易，只想赶快毕业，但在离开之时，才发现这段时光的美好与珍贵，这是人生中的宝贵财富，你会惊讶于自己的坚韧，更会感谢自己的付出。这段时光，也离不开大家的指导与支持，特别感谢李东老师的谆谆教诲，感谢各位老师的无私帮助，感谢统计中心提供的广阔平台，感谢各位小伙伴的鼓励与陪伴。衷心祝愿统计中心越办越好，将统计学发扬光大。

余博

导师：侯琳 / 毕业去向：福瑞德哈金森癌症研究中心博士后



中心越办越好！

统计中心的各位老师通过不懈的努力为我们提供了丰富的学习、研究资源与自由活泼的学术氛围。在中心各位老师的培养下，我在理论与应用、科研与学术交流等诸多方面得到了持续的进步。非常感谢中心的各位老师！感谢侯老师不断的引导与鼓励我。在未来，我将更加坚定自己的选择与追求，继续努力，做自己想做的研究，不辜负中心的培养。同时也感谢中心的各位同学，和你们相处的日子真的很开心。祝各位同学学业有成；祝中心的各位老师工作顺利、生活幸福；祝中心越办越好！

罗声旋

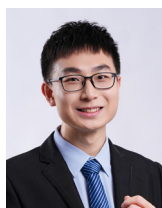
导师：俞声 / 毕业去向：华为公司



感谢博士研究生期间统计学研究中心在我的学术、学习和生活等全方位的培养和照顾。统计中心不仅关注科研能力的培养，也给了我很多接触与统计相关的真实数据和项目的机会。在统计中心的这段时间，我的数据分析能力，从零开始研究一个问题的能力，以及用所学知识解决实际问题的能力都得到了提升。经过这么多年的发展，统计中心的规模和科研成果都有了显著的提升，在此希望统计中心越来越好！师弟师妹们毕业顺利！

朱珂

导师：刘汉中 / 毕业去向：北卡罗来纳州立大学博士后



在清华大学统计学研究中心度过的岁月为我的学术之路打下了坚实的基础。刘汉中老师的悉心指导和全力支持，帮助我克服了读博路上的重重困难，让我感受到专注学术探索的魅力。读博阶段经历的疫情，也让我意识到坚持运动、保持身心健康的重要性。统计学研究中心浓厚的学术氛围以及所有老师和同学们的热心支持，让我在求学过程中倍感温暖和鼓励，谢谢大家！未来，我将怀揣对统计学的热爱继续前行。也祝愿统计中心蒸蒸日上，培养出更多优秀的统计学人才。



-2023 清华大学统计学研究中心优秀大学生夏令营 -

2023 年 6 月底，清华大学统计学研究中心优秀大学生夏令营活动成功举办。本届夏令营共收到来自 101 所高校的 308 名同学的申请，经过层层筛选，来自清华大学、北京大学、北京师范大学、中国人民大学、中国科学技术大学、南开大学等多所国内知名高校的共 30 名同学成功入营。





社会服务及影响

- 社会服务 -

杨立坚

- 2023 年 06 月 哈尔滨工业大学夏季学期学位核心课程“非参数统计”授课教师
- 2023 年 04 月 第二十四届京津冀青年概率统计研讨会组织委员会委员
- 2023 年 03 月—2023 年 7 月 程序委员会委员，首届几何与统计研讨会 Member, Program Committee, The First Symposium of Geometry and Statistics
- 2022 年 12 月—2023 年 7 月 程序委员会委员，全国统计与数据科学联合会议 Member, Program Committee, Joint Conference on Statistics and Data Science in China
- 2022 年 07 月—至今 清华大学理科学术委员会委员
- 2022 年 07 月—至今 科学出版社统计与数据科学系列丛书编委
- 2022 年 07 月—至今 苏州市现场统计研究会第六届监事会顾问
- 2022 年 07 月—至今 中国现场统计研究会生存分析分会第四届理事会常务理事
- 2022 年 07 月—至今 中国现场统计研究会高维数据分会第一届理事会常务理事
- 2021 年 04 月—至今 亚太地区概率统计讨论班理事 (Asia-Pacific Seminar in Probability and Statistics APSPS Board Member)
- 2019 年 10 月—至今 清华大学第十二届学位评定数学分委员会委员
- 2019 年 09 月—至今 中国指挥与控制学会医工结合专业委员会委员

邓 柯

- 2023 年 04 月 ICSCA 2023 International Conference in Chengdu 程序委员会委员
- 2022 年 04 月—至今 International Conference on Bioinformatics and Biomedicine 程序委员会委员
- 2021 年 12 月—2026 年 11 月 清华大学第九届教职工代表大会代表
- 2021 年 12 月—2024 年 12 月 清华大学质量与可靠性研究院副院长
- 2021 年 09 月—2024 年 09 月 清华大学求真书院学生选课指导委员会委员
- 2021 年 01 月—2023 年 12 月 国家抗肿瘤药物临床应用专家委员会委员
- 2020 年 01 月—2022 年 12 月 清华大学教学顾问组成员
- 2020 年 01 月—2022 年 12 月 清华大学信息化用户代表委员会委员
- 2019 年 04 月—至今 北京智源人工智能研究院研究员
- 2018 年 12 月—2022 年 12 月 北京生物医学统计与数据管理研究会副会长
- 2018 年 12 月—2022 年 12 月 中国青年统计学家协会副会长
- 2018 年 12 月—2022 年 12 月 北京市统计学会理事
- 2018 年 10 月—2023 年 09 月 高等教育出版社第二届“现代统计学系列丛书”编委会成员
- 2018 年 01 月—2021 年 12 月 Board Member of IASC-ARS (国际计算统计学会亚太地区分会理事)
- 2017 年 11 月—2021 年 11 月 中国现场统计研究会理事
- 2017 年 04 月—2025 年 10 月 中国现场统计研究会环境与资源分会常务理事
- 2017 年 03 月—2022 年 02 月 中国研究型医院学会医疗信息分会医疗和临床科研大数据应用专委会委员
- 2014 年 10 月—2022 年 10 月 中国数学会概率统计学会理事

李 东

- 2021年09月－2026年09月 北京应用统计学会理事
- 2020年12月－2024年12月 全国工业统计学教学研究会数字经济与区块链技术协会常务理事
- 2019年10月－2023年10月 北京大数据协会常务理事
- 2019年04月－2023年04月 中国青年统计学家协会常务理事
- 2018年12月－2022年12月 全国工业统计学教学研究会第9届理事会常务理事
- 2018年10月－2022年10月 中国概率统计学会第11届理事会副秘书长
- 2017年03月－2021年03月 中国现场统计研究会计算统计分会理事

侯 琳

- 2017年－至今 中国现场统计研究会计算统计分会 秘书长、常务理事
- 2020年－至今 中国自动化学会智能健康与生物信息专委会 委员
- 2021年－至今 中国计算机学会生物信息专委会 通讯委员

俞 声

- 2022年05月 KDD 2022 Health Day Program Committee Member
- 2020年12月－2024年12月 中国现场统计研究会中国旅游大数据分会理事
- 2019年05月－至今 全国工业统计学教学研究会中国青年统计学家协会理事
- 2019年03月－至今 中国现场统计研究会数据科学与人工智能分会理事
- 2017年03月－至今 北京生物医学统计与数据管理研究会副秘书长
- 2017年03月－至今 中国现场统计研究会计算统计分会理事

刘汉中

- 2022年－2026年 全国工业统计学教学研究会理事
- 2021年－2026年 北京应用统计学会理事
- 2019年－2023年 全国工业统计学教学研究会青年统计学家协会理事

林 乾

- 2019年04月－至今 北京智源人工智能研究院“青年科学家”
- 2018年11月－至今 中国现场统计研究会计算统计分会副秘书长

吴未迟

- 2020年12月－2024年12月 全国工业统计学教学研究会数字经济与区块链技术协会理事

杨朋昆

- 2023年 Conference on Learning Theory (COLT) 2023年会议委员会成员

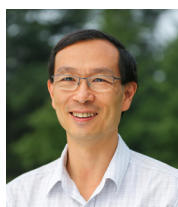
邓婉璐

- 2022年12月－2026年12月 中国现场统计研究会因果推断分会理事
- 2023年04月－2027年04月 中国现场统计研究会教育统计与管理分会理事



- 学术杂志服务 -

统计学研究中心师资团队均在各自领域有很深厚的学术造诣和学术地位，常年为数理统计领域、计量经济学领域、医学信息学领域、生物医学领域、数据科学领域等国际知名学术期刊审稿，如 *The Annals of Statistics*、*Journal of the American Statistical Association*、*Journal of the Royal Statistical Society: Series B*、*Biometrika*、*Journal of Econometrics*、*Journal of Time Series Analysis*、*Statistica Sinica*、*PLOS*、*Genetics*、*Briefings in Bioinformatics*、*Journal of Machine Learning Research*、*Bernoulli*、*BMC Bioinformatics*、*IEEE Transactions on Knowledge and Data Engineering*、*IEEE Transactions on Big Data*、*The Annals of Applied Statistics*、*Biostatistics & Epidemiology* 等。



杨立坚 教授

- *Stat*
- *Statistica Sinica*
- *Brain Sciences*

副主编

副主编

编委



邓柯 副教授

- 《统计与精算》
- 《应用数学与力学》
- *Statistica Sinica*
- 《数字人文》
- 《应用概率统计》

编委 (2022 年 01 月—2025 年 12 月)

编委 (2020 年 11 月—2023 年 11 月)

副主编 (2021 年 08 月—2025 年 07 月)

编委 (2019 年 11 月—2023 年 10 月)

编委 (2019 年 05 月—2023 年 04 月)



侯琳 副教授

- *Statistics in Biosciences*
- *Quantitative Biology*

编委

编委



编辑：侯琳
执行编辑：侯禹珊

清华大学统计学研究中心

地址：北京市海淀区清华大学伟清楼 212 室 (100084)

电话：010-62786091 传真：010-62783842

网址：www.stat.tsinghua.edu.cn

Center for Statistical Science of Tsinghua University

Address: Weiqing Building 212, Tsinghua University, Beijing 100084, China

Tel: +86 10-62786091 Fax: +86 10-62783842

[Http://www.stat.tsinghua.edu.cn](http://www.stat.tsinghua.edu.cn)



扫码关注微信公众号
水木数据派