清華大學统计学研究中心

Center for Statistical Science, Tsinghua University

年度报告

Annual Report

2021.07-2022.06

July 2021 to June 2022

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
- 行政团队

3 学科团队
- 中心教员
- 博士后
- 博士研究生

4 学科发展与人才引进
- 新晋升教员
- 合作研究

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
- 统计咨询服务

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 7
- 本科生培养
- 研究生培养
- 优秀大学生夏令营

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，行政事务挂靠工业工程系。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 8 人、助理教授 4 人、讲师 2 人、博士研究生 48 人。（截至 2022 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。中心还组织专业力量创立统计咨询中心，为清华师生和社会各界提供高质量的统计咨询和数据分析服务。

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；在 QS 国际学科排名中，清华大学“统计与运筹学”由第 26 名（2016 年）提升至 16 名（2020 年、2021 年），并保持稳定，接近国际一流学科水平；近两年泰晤士中国学科评级均为 A +。

2015

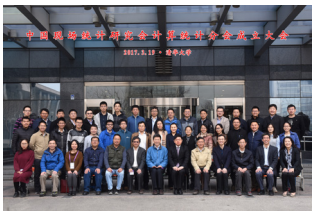
- 6月，中心成立仪式
- 9月，首届博士研究生入学

2016

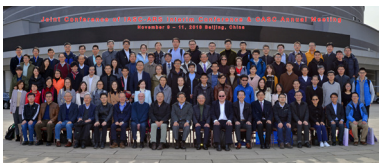
- 9月，开设本科生“统计学辅修”项目

2017

- 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



2018



- 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”

2019

- 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”

2020

- 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”

2021

- 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

组织架构

学术委员会

刘 军　主任

- 哈佛大学统计系
  终身教授

陈 嵘　委员

- 罗格斯大学统计系
  终身教授

沈晓彤　委员

- 明尼苏达大学统计学院
  终身教授

杨立坚　委员

- 清华大学统计学研究中心
  长聘教授

顾问委员会

耿 直　教授

- 北京大学数学科学
  学院教授

黎子良　教授

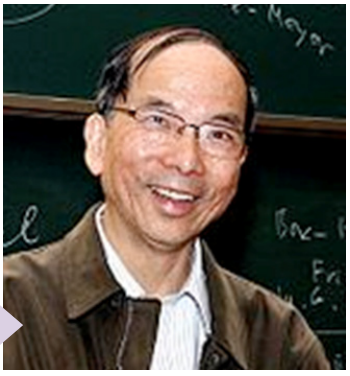
- 斯坦福大学统计系
  终身教授

Prof. Terry Speed

- 加州大学伯克利分校统
  计系和 WEHI终身教授
- 澳洲科学院院士

吴建福　教授

- 佐治亚理工大学工业与系
  统工程学院终身教授
- 美国工程院院士



行政团队

邓 柯
执行主任

田 园
办公室主任

侯禹珊
宣传主管

李志忠　委员

- 清华大学工业工程系长聘教授

林希虹　委员

- 哈佛大学生物统计系终身教授
- 美国国家医学院院士

赵宏宇　委员

- 耶鲁大学公共卫生学院生物统计系终身教授

朱　宇　委员

- 普渡大学统计系终身教授

Prof. Donald Rubin

- 清华大学特聘教授
- 美国科学院院士

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

王永雄　教授

- 斯坦福大学统计系终身教授
- 美国科学院院士

马 腾
行政助理

王 泽
行政助理

宋希婷
科研助理

学科团队

- 中心教员 -

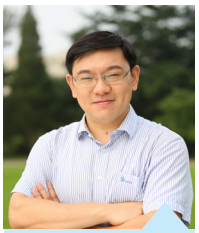
杨立坚 教授

- 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
- 美国统计协会会士（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

- 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



邓 柯 副教授

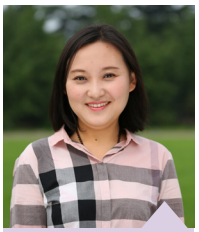
李 东 副教授

- 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：金融计量经济学、非线性时间序列分析、网络与大数据

- 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

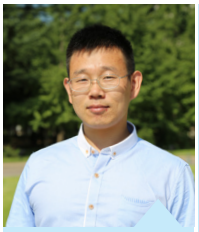


侯 琳 副教授

- 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

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

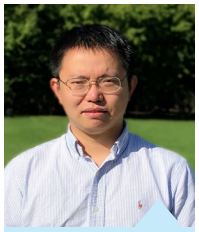
俞 声 副教授



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联合培养博士
- 加州大学伯克利分校统计系博士后

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

吴未迟 副教授

- 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

王天颖 助理教授

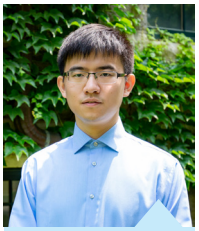
- 德克萨斯 A&M 大学统计学博士
- 哥伦比亚大学生物统计系博士后

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

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



杨朋昆 助理教授

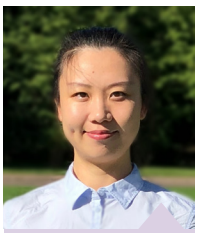
- 伊利诺伊大学香槟分校电子与计算机工程博士
- 普林斯顿大学电子工程系博士后
- 国家级人才计划入选者

研究方向：高维统计理论、机器学习、算法及优化

胡志睿 助理教授

- 哈佛大学统计学博士

研究方向：贝叶斯统计、统计计算、系统发生学、生物信息学



邓婉璐 讲师

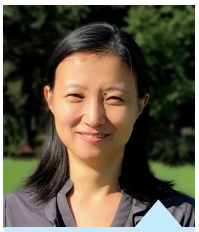
- 北京大学统计学博士
- 宾夕法尼亚大学联合培养博士
- 复旦大学上海数学中心博士后

周在莹 教学副教授

- 清华大学统计学博士

王江典 讲师

- 北卡罗来纳州立大学统计学博士



高梦昭 高级咨询师

- 北京交通大学管理学博士
- 宾夕法尼亚州立大学应用统计硕士

- 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

陈 锐

- 合作导师：林乾
- 清华大学数学专业理学博士
- 2021 年 7 月加入中心

研究方向：充分性降维问题、神经网络

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

- 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、王海沣、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　沄、王羽超、王梓涵、易盈淮、于丁一、张皓博、赵政昀

学科发展与人才引进

- 新晋升教员 -

周在莹 晋升副教授

•2022年清华大学课程思政示范课程、示范教师
•2021年清华大学年度教学优秀奖

- 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**, Jiangyan Wang and **Lijian Yang*** (2022). Kolmogorov-Smirnov simultaneous confidence bands for time series distribution function. *Computational Statistics* **37** (3): 1015-1039.

◆ **Zening Song** and **Lijian Yang*** (2022). Statistical inference for ARMA time series with moving average trend. *Journal of Nonparametric Statistics* **34** (2): 357-376.

◆ **Jiakun Jiang**, Li Cai and **Lijian Yang*** (2022). Simultaneous confidence band for the difference of regression functions of two samples. *Communications in Statistics-Theory and Methods* **51** (11): 3556-3572.

◆ Yan Fang, Lan Xue*, Carlos Martins-Filho and **Lijian Yang** (2022). Robust estimation of additive boundaries with quantile regression and shape constraints. *Journal of Business and Economic Statistics* **40** (2): 615-628.

◆ Jiangyan Wang, Lijie Gu and **Lijian Yang*** (2022). Oracle-efficient estimation for functional data error distribution with simultaneous confidence band. *Computational Statistics and Data Analysis* **167**: 107363.

◆ **Kun Huang**, Dian Chen, Fei Wang* and **Lijian Yang*** (2021). Prediction of dispositional dialectical thinking from resting-state electroencephalography. *Brain and Behavior* **11** (9)：e2327.

◆ Pei Cao, Lei Zhang, Yang Yang, Xiaodan Wang, Zhaoping Liu, Jianwen Li, Liyuan Wang, Sookja Chung, **Moqin Zhou**, **Ke Deng**, Pingping Zhou and Pinggu Wu (2022). Analysis of furan and its major furan derivatives in coffee products on the Chinese market using HS-GC-MS and the estimated exposure of the Chinese population. *Food Chemistry* **387**: 132823.

◆ Hao Xu#, **Tingxuan Han#**, **Haifeng Wang**, Shanggui Liu, Guanghao Hou, Guanchao Jiang, Jian Zhou* and **Ke Deng*** (2022). Detection of blood stains using computer vision-based algorithms and their association with postoperative outcomes in thoracoscopic lobectomies. *European Journal of Cardio-Thoracic Surgery*: ezac154.

◆ **Yichao Li#**, Wenshuo Wang#, **Ke Deng*** and **Jun S. Liu*** (2022). Stratification and optimal resampling for sequential Monte Carlo. *Biometrika* **109** (1): 181–194.

◆ **Changzai Pan**, Maosong Sun, **Ke Deng*** (2022). TopWORDS-Seg: simultaneous text segmentation and word discovery for open-domain Chinese texts via Bayesian Inference. *In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Dublin, Ireland. Association for Computational Linguistics*: 158–169.

◆ **Yingkai Jiang**, Xinshu Zhao, Lixing Zhu, **Jun S**. **Liu** and **Ke Deng*** (2021). Total- effect test is superfluous for establishing complementary mediation. *Statistica Sinica* **31** (3): 1961-1983.

◆ Lingyu Sun and **Dong Li*** (2021). Change-point detection for expected shortfall in time series. *Journal of Management Science and Engineering* **6** (3): 324-335.

◆ **Feiyu Jiang**, **Dong Li** and Ke Zhu* (2021). Adaptive inference for a semiparametric GARCH model. *Journal of Econometrics* **224** (2): 306-329.

◆ **Xinyu Zhang*** and **Howell Tong** (2021). Asymptotic theory of principal component analysis for time series data with cautionary comments. *Journal of the Royal Statistical Society: Series A* **185** (2): 543-565.

◆ Jialin Liu, Yiling Li, **Dong Li**, Yibaina Wang and Sheng Wei (2022). The burden of coronary heart disease and stroke attributable to dietary cadmium exposure in Chinese adults, 2017. *Science of the Total Environment* **825**: 153997.

◆ **Shuang Song**, Wei Jiang, Yiliang Zhang, **Lin Hou** and Hongyu Zhao* (2022). Leveraging LD eigenvalue regression to improve the estimation of SNP heritability and confounding inflation. *The American Journal of Human Genetics* **109**: 802-811.

◆ Xiaoyang Chen, Shengquan Chen, **Shuang Song**, Zijing Gao, **Lin Hou**, Xuegong Zhang, Hairong Lv and Rui Jiang* (2022). Cell type annotation of single-cell chromatin accessibility data via supervised Bayesian

embedding. *Nature Machine Intelligence* **4** (2): 116-126.

◆ **Shuang Song**, **Lin Hou*** and **Jun S. Liu*** (2022). A data-adaptive Bayesian regression approach for polygenic risk prediction. *Bioinformatics* **38** (7): 1938-1946.

◆ **Hanmin Guo**, **Lin Hou** and Yu Zhu* (2022). Minimal-field for flexible sufficient dimension reduction. *Electronic Journal of Statistics* **16**:1997-2032.

◆ Bingxiang Xu, Mingjie Lu, Linlin Yan, Minghui Ge, Yong Ren, Ru Wang, Yongqian Shu, **Lin Hou*** and Hao Guo* (2021). A pan-cancer analysis of predictive methylation signatures of response to cancer immunotherapy. *Frontiers in Immunology* **12**: 796647.

◆ **Nayang Shan**, Yuhan Xie, **Shuang Song**, Wei Jiang, Zuoheng Wang* and **Lin Hou*** (2021). A novel transcriptional risk score for risk prediction of complex human diseases. *Genetic Epidemiology* **45** (8):811-820.

◆ **Shuang Song**, **Nayang Shan**, Geng Wang, Xiting Yan, **Jun S. Liu** and **Lin Hou*** (2021). Openness weighted association studies: leveraging personal genome information to prioritize non-coding variants. *Bioinformatics* **37** (24):4737-4743.

◆ **Zheng Yuan**, **Zhengyun Zhao**, Haixia Sun, Jiao Li, Fei Wang and **Sheng Yu*** (2022). CODER: knowledge-infused cross-lingual medical term embedding for term normalization. *Journal of Biomedical Informatics* **126**:103983.

◆ Sihang Zeng, **Zheng Yuan** and **Sheng Yu*** (2022). Automatic biomedical term clustering by learning fine-grained term representations. *In Proceedings of the 21st Workshop on Biomedical Language Processing, Dublin, Ireland. Association for Computational Linguistics*: 91–96.

◆ **Hongyi Yuan**, **Zheng Yuan**, Ruyi Gan, Jiaxing Zhang, Yutao Xie and **Sheng Yu*** (2022). BioBART: pretraining and evaluation of a biomedical generative language model. *In Proceedings of the 21st Workshop on Biomedical Language Processing, Dublin, Ireland. Association for Computational Linguistics*: 97–109.

◆ **Shengxuan Luo** and **Sheng Yu*** (2022). An accurate unsupervised method for joint entity alignment and dangling entity detection. *In Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland. Association for Computational Linguistics*: 2330–2339.

◆ Ningyu Zhang, Mosha Chen, Zhen Bi, Xiaozhuan Liang, Lei Li, Xin Shang, Kangping Yin, Chuanqi Tan, Jian Xu, Fei Huang, Luo Si, Yuan Ni, Guotong Xie, Zhifang Sui, Baobao Chang, Hui Zong, **Zheng Yuan**, Linfeng Li, Jun Yan, Hongying Zan, Kunli Zhang, Budhou Tang and Qingcai Chen (2022). CBLUE: a Chinese biomedical language understanding evaluation benchmark. *In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Dublin, Ireland. Association for Computational Linguistics*: 7888–7915.

◆ **Zheng Yuan**, Chuanqi Tan, and Songfang Huang (2022). Code synonyms do matter: multiple synonyms matching network for automatic ICD coding. *In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Dublin, Ireland. Association for Computational Linguistics*: 808–814.

◆ **Zheng Yuan**, Chuanqi Tan, Songfang Huang and Fei Huang (2022). Fusing heterogeneous factors with triaffine mechanism for nested named entity recognition. *In Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland. Association for Computational Linguistics*: 3174–3186.

◆ **Qian Lin**，Zhigen Zhao and **Jun S. Liu** (2021). Testing model utility for single index models under high dimension. *Festschrift in Honor of R. Dennis Cook*：65-86.

◆ Holger Dette and **Weichi Wu*** (2022). Prediction in locally stationary time series. *Journal of Business & Economic Statistics* **40** (1): 370-381.

◆ **Tianying Wang**, Wodan Ling, Anna M Plantinga, Michael C Wu and Xiang Zhan (2022). Testing microbiome association using integrated quantile regression models. *Bioinformatics* **38** (2): 419-425.

◆ **Jingyi Zhang**, Wenxuan Zhong and Ping Ma* (2021). A review on modern computational optimal transport methods with applications in biomedical research. *Modern Statistical Methods for Health Research* **1**: 279-300.

◆ Cong Fang, Jason Lee, **Pengkun Yang** and Tong Zhang (2021). Modeling from features: a mean-field framework for over-parameterized deep neural networks. *Proceedings of Thirty Fourth Conference on Learning Theory (COLT)*: 1887–1936.

◆ Tianyang Wang, Xingjian Li, **Pengkun Yang**, Guosheng Hu, Xiangrui Zeng, Siyu Huang, Cheng-Zhong Xu and Min Xu (2022). Boosting active learning via improving test performance. *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI).*

◆ Huachen Zhang, **Ke Zhu**, **Jiangdian Wang**, Xiaoli Lv* (2022). The use of a new classification in endovascular treatment of dural arteriovenous fistulas. *Neuroscience Informatics* **2** (2):100047.

◆ **王江典** ,**沈翀** ,杨蕾 ;**高梦昭** ,王红 ,**邓柯** (2022). 线上和线下本科教学质量的比较分析——基于清华大学教学评估数据 .*《中国电化教育》* **3**: 90-95.

In Press

◆ **Kun Huang**, **Sijie Zheng** and **Lijian Yang*** (2022+). Inference for dependent error functional data with application to event-related potentials. *TEST*.

◆ **Chen Zhong** and **Lijian Yang*** (2022+). Statistical inference for functional time series: autocovariance function. *Statistica Sinica*.

◆ **Jie Li** and **Lijian Yang*** (2022+). Statistical inference for functional time series. *Statistica Sinica*.

◆ Yang Yang and **Ke Deng*** (2022+). Generalized theme dictionary models for association pattern discovery. *The Annals of Applied Statistics*.

◆ Wanchuang Zhu, Yingkai Jiang, **Jun S. Liu** and **Ke Deng*** (2022+). Partition-mallows model and its inference for rank aggregation. *Journal of the American Statistical Association*.

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

◆ Donghang Luo, Ke Zhu, Huan Gong and **Dong Li*** (2022+). Testing error distribution by kernelized Stein discrepancy in multivariate time series models. *Journal of Business & Economic Statistics*.

◆ Feiyu Jiang, **Dong Li**, Wai Keung Li and Ke Zhu (2022+). Testing and modelling for the structural change in covariance matrix time series with multiplicative form. *Statistica Sinica.*

◆ **Hanmin Guo**, **Lin Hou**, Yu Shi , Sheng Chih Jin, Xue Zeng, Boyang Li, Richard Lifton, Martina Brueckner, Hongyu Zhao* and Qiongshi Lu* (2022+). Quantifying concordant genetic effects of de novo mutations on multiple disorders. *eLife*.

◆ **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*.

◆ **Shengxuan Luo**, Huaiyuan Ying and **Sheng Yu*** (2022+). Label refinement via contrastive learning for

distantly-supervised named entity recognition. *Findings of NAACL 2022*.

◆ **Hongyi Yuan**, **Zheng Yuan** and **Sheng Yu*** (2022+). Generative biomedical entity disambiguation via knowledge base-guided pre-training and synonyms-aware fine-tuning. *NAACL 2022*.

◆ Qiao Jin, **Zheng Yuan**, Guangzhi Xiong, Qianlan Yu, Huaiyuan Ying, Chuanqi Tan, Mosha Chen, Songfang Huang, Xiaozhong Liu and **Sheng Yu*** (2022+). Biomedical question answering: a survey of approaches and challenges. *ACM Computing Surveys*.

◆ Xinhe Wang, Tingyu Wang and **Hanzhong Liu** (2022+). Rerandomization in stratified randomized experiments. *Journal of the American Statistical Association*.

◆ **Hanzhong Liu**, Fuyi Tu and Wei Ma (2022+). Lasso-adjusted treatment effect estimation under covariate-adaptive randomization. *Biometrika*.

◆ **Ke Zhu** and **Hanzhong Liu** (2022+). Confidence intervals for parameters in high-dimensional sparse vector autoregression. *Computational Statistics and Data Analysis*.

◆ **Hanzhong Liu**, **Jiyang Ren** and Yuehan Yang (2022+). Randomization-based joint central limit theorem and efficient covariate adjustment in randomized block 2^K factorial experiments. *Journal of the American Statistical Association*.

◆ Subhra Dhar and **Weichi Wu** (2022+). Comparing time varying regression quantiles under shift invariance. *Bernoulli*.

◆ **Tianying Wang***, Iuliana Ionita-Laza and Ying Wei (2022+). Integrated Quantile RAnk Test (iQRAT) for gene-level associations. *The Annals of Applied Statistics*.

◆ Lauren C Houghton, Ying Wei, **Tianying Wang**, Mandy Goldberg, Alejandra Paniagua-Avila, Rachel L Sweeden, Angela Bradbury, Mary Daly, Lisa A Schwartz, Theresa Keegan, Esther M John, Julia A Knight, Irene L Andrulis, Saundra S Buys, Caren J Frost, Karen O' Toole, Melissa L White, Wendy K Chung and Mary Beth Terry (2022+). Body mass index rebound and pubertal timing in girls with and without a family history of breast cancer: the LEGACY girls study. *International Journal of Epidemiology* .

◆ **Jingyi Zhang**, Cheng Meng, Jun Yu, Mengrui Zhang, Wenxuan Zhong and Ping Ma* (2022+). An optimal transport approach for selecting a representative subsample with application in efficient kernel density estimation. *Journal of Computational and Graphical Statistics*.

◆ **Jingyi Zhang**, Ping Ma, Wenxuan Zhong and Cheng Meng* (2022+). Projection‐based techniques for high-dimensional optimal transport problems. *Wiley Interdisciplinary Reviews: Computational Statistics*.

◆ **Jingyi Zhang**, Huolan Zhu, Chenguang Yang, Yongkai Chen, Huimin Cheng, Yi Li, Fang Wang* and Wenxuan Zhong* (2022+). Ensemble machine learning approach for screening of coronary heart disease based on echocardiography and risk factors. *BMC Medical Informatics and Decision Making*.

◆ Natalie Doss, **Pengkun Yang**, Yihong Wu and Harrison Zhou (2022+). Optimal estimation of high-dimensional Gaussian mixtures. *The Annals of Statistics*.

书籍编写：

谷成明，李一，王斌辉，**王江典**等。《真实世界数据与证据：引领研究规范，赋能临床实践》科学技术文献出版社
出版时间：2022年 5月

2021
07

邓柯课题组在 AOAS发表论文提出电磁超材料快速设计新方法

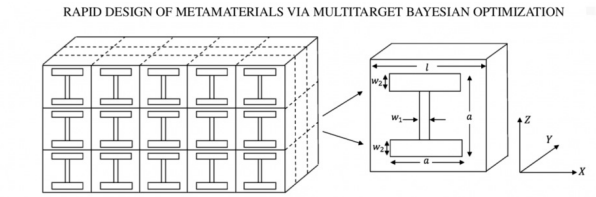
The Annals of Applied Statistics
2021, Vol. 15, No. 2, 768–796
https://doi.org/10.1214/20-AOAS1426
© Institute of Mathematical Statistics, 2021

RAPID DESIGN OF METAMATERIALS VIA MULTITARGET BAYESIAN OPTIMIZATION

BY YANG YANG[1], CHUNLIN JI[2] AND KE DENG[3]

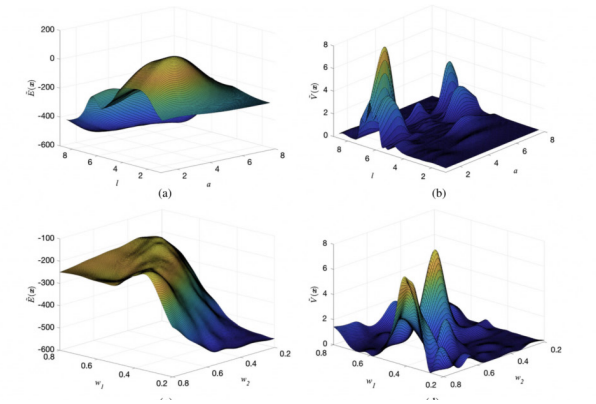
[1]Department of Mathematical Sciences & Center for Statistical Science, Tsinghua University, yyang15@mails.tsinghua.edu.cn
[2]Kuang-Chi Institute of Advanced Technology, Shenzhen, China, chunlin.ji@kuang-chi.com
[3]Center for Statistical Science and Department of Industrial Engineering, Tsinghua University, kdeng@tsinghua.edu.cn

杨洋博士 第一作者　　季春霖博士 共同通讯作者　　邓柯副教授 通讯作者



RAPID DESIGN OF METAMATERIALS VIA MULTITARGET BAYESIAN OPTIMIZATION

FIG. 1. *A schematic diagram showing the relationship between the macrostructure and the microstructure of a metamaterial prototype.*



我中心邓柯副教授课题组在应用统计国际顶尖期刊The Annals of Applied Statistics (AOAS)发表题为“Rapid Design of Metamaterials via Multitarget Bayesian Optimization”的研究论文，提出了电磁超材料快速设计的新方法。曾在邓柯课题组攻读博士学位的杨洋博士（清华大学 2015级博士生）是该文的第一作者，邓柯副教授与深圳光启高等理工研究院副院长季春霖博士作为论文的共同通讯作者联合指导了相关研究和论文撰写。

电磁超材料由大量人工设计建造的结构单元构成，能呈现出光学隐身、聚焦等自然材料所不具备的超常电磁性能，在许多领域具有重要应用，在学术界和工业界都引起了广泛的关注。然而，传统的电磁超材料设计一般是借助不断试错的探索性实验进行的，缺乏有理论保证的系统方法，需要耗费大量时间和计算资源，严重制约了电磁超材料的实际应用。

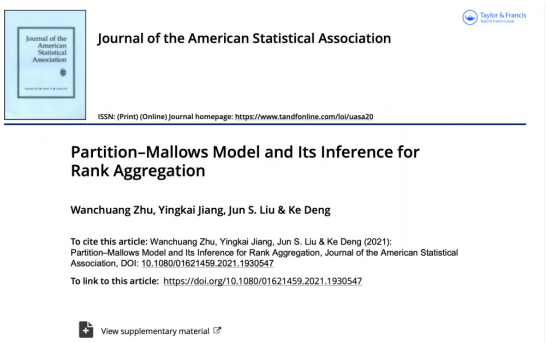
该论文将电磁超材料设计问题凝练为一个多目标优化问题，提出了一种基于贝叶斯优化的协同设计方法，大幅提高了设计效率，实现了电磁超材料的快速设计。该方法首先根据实际问题的背景，将一个结构单元的无穷维响应曲线降维成两个简单响应（即响应曲线的均值和方差），并用高斯过程对这两个简单响应进行建模；进而，利用这些简单响应的统计模型构建起便捷的代理模型，对不同结构单元的响应给出定量预测，并同步考虑复杂电磁超材料所涉及的多个设计目标来构造联合采集函数，在贝叶斯优化的框架下选择最有利的实验设计路径。论文从理论上证明了该方法的有效性，并通过计算机模拟实验验证了该方法具有远超已有方法的计算和搜索效率。

从理论的角度看，该论文首次为复杂电磁超材料的设计建立了具有理论保证的系统方法；从应用的角度看，该论文提出的方法对复杂电磁超材料的设计效率有了数量级上的提升。相关方法的应用不仅局限于电磁超材料设计领域，还有潜力拓展到许多具有类似问题的应用场景。该项工作是统计学习和材料工程两个领域交叉融合的成果，是重要应用问题驱动交叉学科研究的一个成功实例。

该研究工作获得国家自然科学基金（Grant 11931001 & 11771242）、北京智源人工智能研究院（Grant BAAI2019ZD0103）、超材料电磁调制技术国家重点实验室和广东省超材料微波射频重点实验室的资助。

2021
08

邓柯课题组在 JASA发表论文提出排名聚合新方法


Journal of the American Statistical Association

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

Partition–Mallows Model and Its Inference for Rank Aggregation

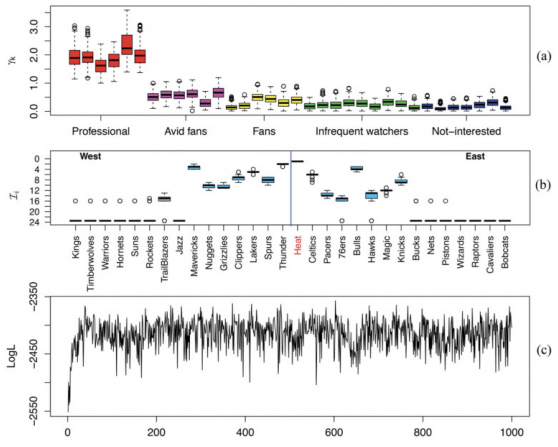
Wanchuang Zhu, Yingkai Jiang, Jun S. Liu & Ke Deng

To cite this article: Wanchuang Zhu, Yingkai Jiang, Jun S. Liu & Ke Deng (2021): Partition–Mallows Model and Its Inference for Rank Aggregation, Journal of the American Statistical Association, DOI: 10.1080/01621459.2021.1930547

To link to this article: https://doi.org/10.1080/01621459.2021.1930547

View supplementary material


朱万闯博士 第一作者　姜瑛恺博士 共同作者　刘军教授 共同作者　邓柯副教授 通讯作者



我中心邓柯副教授课题组在统计国际顶尖期刊 Journal of the American Statistical Association (JASA) 发表题为“Partition-Mallows Model and Its Inference for Rank Aggregation”的研究论文，提出了一种推断排名聚合的新方法。曾在邓柯课题组工作的朱万闯博士是该文的第一作者，姜瑛恺博士和刘军教授为共同作者，邓柯副教授是论文的通讯作者。

排名聚合是指如何聚合从不同信息源获得的关于某些个体的排序，从而得到一个更加“精确”的排序。例如，有 m位评委为 n名运动员的能力进行排序。排名聚合致力于对这 m个排序进行整合分析以得到一个新的排序，能够更加准确地反映 n名运动员能力的高低。现实中，m位评委的可靠性可能会存在差异，部分可靠性较低的评委可能会误导排名聚合的结果。开发基于数据驱动的方法来自动识别不同评委的可靠性，并据此优化排名聚合的结果，具有重要的实际意义。

邓柯和刘军等人曾于 2014年在 JASA发表了题为“Bayesian Aggregation of Order-Based Rank Data”的论文中，提出了一种基于划分模型(partition model) 的排名聚合方法 BARD。BARD将排序对象划分为两个组别，“相关个体组”和“背景个体组”，并假设可靠性高的评委们会以更高的概率将中的个体排位于中的个体之前。该方法能够在有效识别评委可靠性的同时，通过弱化可靠性较差的评委在排名聚合中贡献，来消除他们可能带来的负面作用。但是，该方法简单忽略了和两个组别中各个体的差异，从而在很大程度上损失了组内排名的信息。从应用的角度看，这是该方法的一个重要局限性。

本文在上述工作的基础上，采用更加精细的 Mallows模型对组别的组内排名进行了建模，将 partition模型和 Mallows 模型的优势结合起来，得到了能力更强的排名聚合模型 Partition-Mallows model。该模型构建了对具有复杂结构的排名数据进行定量描述的一般框架，在充分利用和组间及组内的排名信息的基础上，不仅可以有效识别评委可靠性的差异，还能够产出更有效率的排名聚合。我们从理论上证明了该方法的可靠性，并通过大量的计算机模拟和实证研究验证了该方法在处理具有分组结构的排名聚合问题上具有明显优势。

该研究工作获得中国国家自然科学基金 (Grants 11771242 & 11931001)、北京智源人工智能研究院 (Grant BAAI2019ZD0103)和美国国家科学基金 (Grants DMS-1903139 and DMS-1712714)的资助。

2021
09

中心博士研究生斩获国际统计学会 2021年度简· 丁伯根奖一等奖

统计学研究中心 2017级博士研究生李杰，2020级博士研究生胡祺睿斩获国际统计学会 (International Statistical Institute, ISI) 颁发的国际统计学会 2021年度简· 丁伯根奖一等奖(2021 ISI Jan Tinbergen Award Division A First Prize)。

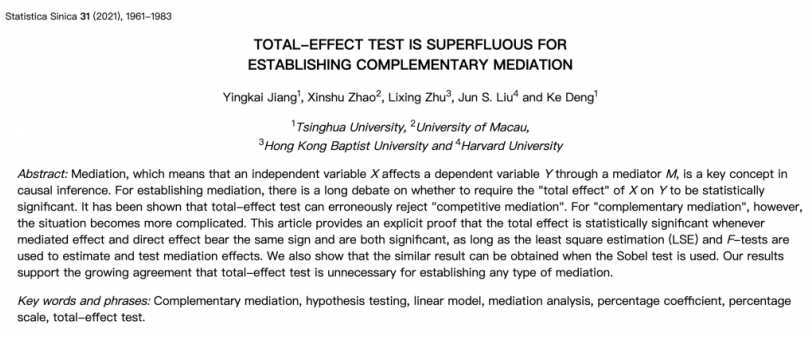
总部在荷兰的国际统计学会(ISI) 是全球三个权威统计学学术组织之一，旨在引领、支持和促进全世界对统计学的理解、发展和良好实践。ISI 颁发的各类荣誉奖项都被国际统计学界高度认可。简· 丁伯根奖命名于获得 1969年首个诺贝尔经济学奖的荷兰学者简· 丁伯根，是从每两年举行一次的世界统计学大会(World Statistics Congress, WSC) 青年统计学者(1987年以后出生) 提交的论文中评选的最佳论文(https://www.isi-web.org/events/isi-awards/tinbergen-award)。其中 Division A的获奖论文必须解决一个对广大发展中国家具有实际意义的应用统计问题。自 2019年开始，获奖者已不再限于发展中国家。自 2013年至 2021年，共有来自多个国家的 14人获奖，其中 3位华人，李杰和胡祺睿是第一次获得一等奖的华人。除此之外，西安电子科技大学数学与统计学院研究生韩路于 2013年获二等奖。

今年共有 3人获奖。李杰和胡祺睿获得 2500欧元奖金，受邀免注册费参加于 7月 11日至 16日在荷兰海牙举行(最终因疫情在线举办) 的国际统计学会第 63届世界统计学大会(The 63rd ISI World Statistics Congress)，并在简· 丁伯根奖会场(Jan Tinbergen Awards Session) 做了 30分钟的邀请报告。

李杰和胡祺睿的获奖论文“非参数回归分析空气污染物浓度的预测区间”(Prediction Interval of Air Pollutants Concentration by Nonparametric Regression Analysis) 将非参数回归模型应用于局部平稳时间序列的趋势，分析了由中国环境监测总站高级工程师张凤英博士提供的西安市 2013年到 2020年间 6种主要空气污染物的每日浓度数据，并构造出了未来 5日各空气污染物浓度的预测区间。论文提出用样条回归(Spline regression) 估计趋势函数，核回归(Kernel regression) 估计方差函数，对所得的近似平稳序列拟合自回归(AR) 模型，再用核分布(Kernel distribution estimator) 方法估计其误差的分位数后，得到了带趋势项自回归时间序列的数据驱动多步向前预测区间。相比于季节性差分整合移动平均自回归(Seasonal ARIMA) 等传统方法产生的预测区间，论文中方法得到的预测区间不仅长度更窄，还具有更好的预测精度和覆盖率。该方法有效解释了空气污染物浓度数据潜在的动态变化规律，并可以精确预测未来五到七日空气污染物的浓度，在污染物管理和早期预防方面有着广泛的应用价值。特别值得一提的是李杰和胡祺睿的获奖论文是在无指导教师直接参与的条件下完成的。

2021
11

邓柯课题组在 Statistica Sinica发表论文证明中介效应分析不需要总效应检验


Statistica Sinica 31 (2021), 1961-1983

TOTAL-EFFECT TEST IS SUPERFLUOUS FOR ESTABLISHING COMPLEMENTARY MEDIATION

Yingkai Jiang[1], Xinshu Zhao[2], Lixing Zhu[3], Jun S. Liu[4] and Ke Deng[1]

[1]Tsinghua University, [2]University of Macau, [3]Hong Kong Baptist University and [4]Harvard University

Abstract: Mediation, which means that an independent variable *X* affects a dependent variable *Y* through a mediator *M*, is a key concept in causal inference. For establishing mediation, there is a long debate on whether to require the "total effect" of *X* on *Y* to be statistically significant. It has been shown that total-effect test can erroneously reject "competitive mediation". For "complementary mediation", however, the situation becomes more complicated. This article provides an explicit proof that the total effect is statistically significant whenever mediated effect and direct effect bear the same sign and are both significant, as long as the least square estimation (LSE) and *F*-tests are used to estimate and test mediation effects. We also show that the similar result can be obtained when the Sobel test is used. Our results support the growing agreement that total-effect test is unnecessary for establishing any type of mediation.

Key words and phrases: Complementary mediation, hypothesis testing, linear model, mediation analysis, percentage coefficient, percentage scale, total-effect test.

我中心邓柯副教授课题组在统计学顶尖期刊 Statistica Sinica 发表题为“Total- effect Test is Superfluous for Establishing Complementary Mediation ”的研究论文，从数学上严格地证明了当直接效应和间接效应同方向且均统计显著时，利用最小二乘估计(LSE) 和 F-检验建立中介效应时总效应检验一定是显著的。同时本文还将类似的结果推广到了利用 Sobel检验建立中介效应的场景。曾在邓柯课题组攻读博士学位的姜瑛恺博士(清华大学 2015 级博士生) 是该文的第一作者，邓柯副教授作为通讯作者与澳门大学赵心树教授、香港浸会大学朱力行教授和哈佛大学刘军教授共同指导了该文的研究和撰写。

中介效应模型是因果推断中一类经典的模型，它是指自变量 X 通过中介变量 M 对因变量 Y 产生影响。在社会科学诸多领域的研究中受到研究者的青睐。通常称给定 M的条件下，X 对 Y 的影响为直接效应，X通过 M对 Y产生的影响为间接效应，两者之和为总效应。在建立中介效应时，文献中对于“是否需要 X对 Y的总效应是统计显著的”这一条件是有争议的。已经有研究指出，当直接效应和间接效应符号相反(称为竞争中介) 或直接效应为零(称为完全中介) 时，总效应检验有可能会错误地拒绝中介效应。然而，对于直接效应和间接效应同号(称为互补中介) 的情形，总效应检验的作用并未达成共识。该文创造性地将是否需要对总效应进行检验的问题转化对相关检验拒绝域的包含关系进行几何验证的问题，从而从数学上严格证明了当直接效应和间接效应同方向且均统计显著时，在 LSE-F框架下总效应检验一定显著，在 LSE-Sobel框架下相关结论渐近成立。除上述结论之外，研究团队还利用所构造的几何分析方法，对中介效应的各种情形进行了系统分析，从统计推断和几何分析的双重角度对已有文献中关于中介效应检验的结论给予了新的解读。同时，随机模拟实验的结果与理论结果也是完全契合的。以上这些结论与文献中已有的结果相互印证，支持了一个共同的论断： 在各种情形下建立中介效应都不需要总效应检验。最后，研究团队通过一份社会学研究数据展示利用中介效应模型进行实际数据分析的方法。该研究工作获得国家自然科学基金 (Grants 11771242)、北京智源人工智能研究院 (Grant BAAI2019ZD0103) 的资助。

2022
03

清华统计 x2022ACL:邓柯、俞声两课题组多篇文章被接受

2022年第 60届国际计算语言学协会年会 (Annual Meeting of the Association for Computational Linguistics，简称 ACL)举行，我中心邓柯课题组 18级博士研究生潘长在、俞声课题组 17级博士研究生袁正、18级博士研究生罗声旋几位同学的多篇投稿文章被接受。ACL会议始于 1962年，由国际计算语言学协会主办，是自然语言处理与计算语言学领域最高级别的学术会议。

潘长在同学的论文入选“主会长文”单元，题为“ TopWORDS-Seg:开放域中文文本领域通过贝叶斯推断同时进行文本切词和词语发现的方法 (TopWORDS-Seg: Simultaneous Text Segmentation and Word Discovery for Open-Domain Chinese Texts via Bayesian Inference)”，文章针对于几十年来计算语言学中的一个关键瓶颈，开放域中文文本处理问题展开论述。称之为瓶颈是因为在开放域这种具有挑战性的场景中，文本分词和词语发现经常相互纠缠，且并无可用的训练数据。尚无现有方法可以在开放域中同时实现有效的文本分词和单词发现。该文章通过提出一种基于贝叶斯推理的名为 TopWORDS-Seg 的新方法来填补这一空白，在没有训练语料库和领域词表的情况下具有很好的表现和解释性。该文章通过维基百科数据用一系列实验研究证明了 TopWORDS-Seg 的优势。潘长在是第一作者，邓柯副教授作为通讯作者与清华大学计算机系科学与技术系的孙茂松教授共同指导了该工作。

袁正同学共有三篇文章入选：

入选“主会短文”单元文章：“基于疾病同义词的匹配网络的自动疾病编码(Code Synonyms Do Matter: Multiple Synonyms Matching Network for Automatic ICD Coding)”通过额外利用疾病编码的同义词信息去匹配电子病历中的不同文本以达到更好的自动疾病编码效果，在 MIMIC-3电子病历数据集上得到了超过以往方法的分类效果。

入选“Findings长文”单元文章 ：“使用三仿射融合异质信息的嵌套命名实体识别方法(Fusing Heterogeneous Factors with Triaffine Mechanism for Nested Named Entity Recognition)”通过三仿射变换改进基于片段分类的命名实体识别模型中的片段表示和片段分类方法，在新闻和医疗命名实体识别任务上取得了超过之前的结果。以上两篇文章袁正均为第一作者，与阿里巴巴达摩院团队合作完成。

此外，袁正与浙江大学、鹏程实验室等研究团队合作的论文：“中文医学自然语言处理评测数据集(CBLUE: A Chinese Biomedical Language Understanding Evaluation Benchmark）也入选了“主会长文”单元。

罗声旋同学的论文入选“Findings长文”单元，题为“联合实体对齐和悬空实体识别的高精度无监督方法 (An Accurate Unsupervised Method for Joint Entity Alignment and Dangling Entity Detection) ”，罗声旋为该文的第一作者，其导师俞声副教授为通讯作者。文章针对在对齐两个知识图谱的现实场景中的三个主要问题：(1）不存在等价对应的实体，也即悬空实体，广泛存在于知识图谱中；(2）悬空实体标签和实体对(等价的两个实体）标签难以获得，一个普适的知识图谱对齐方法需要尽可能避免对监督数据的依赖；(3）各对齐之间以及预测对齐与识别悬空实体之间是互相影响的，需要整体地考虑识别悬空实体并对齐等价的实体。该文章首先根据实体的文本语义信息和全局的相似性指导两个知识图谱中的实体嵌入的训练，从而获得实体之间的距离估计。然后，给每个知识图谱添加一个虚拟实体，从而把实体对齐和悬空实体整合为一个统一的最优运输问题，并解这个问题。最终，与虚拟实体对齐的实体为悬空实体，其余对齐为模型预测的等价实体对。一系列实验表明，该文章在不依赖监督数据的情况下，能够达到当前实体对齐任务上的最优表现，并且有高质量的悬空实体识别结果。

2022
04

清华统计 +北大人民：手术影像出血量化助力医学临床实践

我中心邓柯副教授团队与北京大学人民医院胸外科团队合作在胸外科经典期刊《欧洲心胸外科杂志》（European Journal of Cardio-Thoracic Surgery，简称 EJCTS）发表题为“Detection of blood stains using computer vision-based algorithms and their association with postoperative outcomes in thoracoscopic lobectomies”的研究论文，提出了一种基于计算机视觉的术中出血识别与计量方法，并首次将手术视频分析与临床预后关联分析相结合，为手术预后提供指导。

北京大学人民医院博士研究生许昊和中心 21级博士研究生韩庭萱为本文的共同第一作者，中心邓柯副教授和北京大学人民医院的周健副教授、王俊院士为本文的共同通讯作者，中心 19级博士研究生王海洋参与工作。

如今，胸外科手术已逐渐转向微创 (Minimally Invasive Surgery，简称 MIS)范式。与开胸手术相比，视频辅助胸腔镜手术(Video Assisted Thoracoscopic Surgery，简称 VATS）因其较少的术后急性期、较小的肺功能损害、较低的术后发病率以及较短的住院时间，已成为治疗早期和局部晚期非小细胞肺癌患者的标准方式。在 VATS手术期间，数码相机将手术过程拍摄为视频。手术视频中包含大量有关患者术中出血情况的信息，例如出血的时段、冲水操作后的出血状况等。如何通过对手术视频的分析准确有效地量化 VATS视频中的血迹，并提取能够反映手术状态、与预后显著相关的变量是一项尚未解决的挑战。

本文使用北京大学人民医院胸外科 2020年行肺叶切除术的 275例手术视频，利用图像背景信息，提出了一种基于 RGB通道动态阈值的血迹像素识别算法，对血迹像素的判别准则进行实时调整，并在血迹像素识别任务中取得准确性 99.1%，特异性 98.9%，敏感度 99.2%的效果。该方法具有运算简单快速、准确性高的特点，适用于长时间、高清晰度的视频分析。对术中血迹像素进行识别后，本文提取基于血迹像素占比的变量用以刻画术中出血情况，并与患者预后指标进行关联分析，得到了具有统计学显著性的变量，由此形成临床预后方案，为预后管理提供指导。

清华统计 + 国家食品安全风险评估中心：统计学在咖啡中污染物浓度相关结构分析中的应用

2022
04


Food Chemistry 387 (2022) 132823

Contents lists available at ScienceDirect

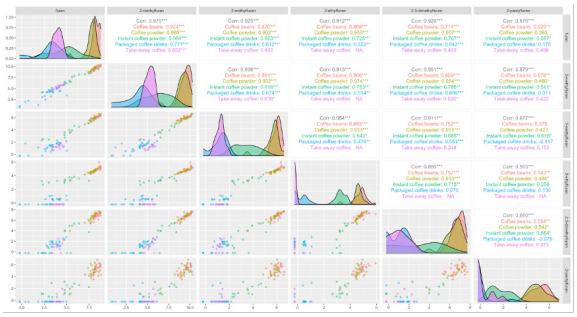
Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Analysis of furan and its major furan derivatives in coffee products on the Chinese market using HS-GC–MS and the estimated exposure of the Chinese population

Pei Cao [a], Lei Zhang [a], Yang Yang [b], Xiao-dan Wang [a], Zhao-ping Liu [a], Jian-wen Li [a], Li-yuan Wang [c], Sookja Chung [d], Moqin Zhou [e], Ke Deng [e], Ping-ping Zhou [b,*], Ping-gu Wu [c,*]

[a] China National Center for Food Safety Risk Assessment, Beijing 100022, China
[b] Chinese Academy of Inspection and Quarantine, Beijing 100176, China
[c] Zhe Jiang Provincial Center for Disease Control and Prevention, Hangzhou 310051, China
[d] Faculty of Medicine, Macau University of Science and Technology, Macao, China
[e] Center for Statistical Science & Department of Industrial Engineering, Tsinghua University, Beijing 100084, China

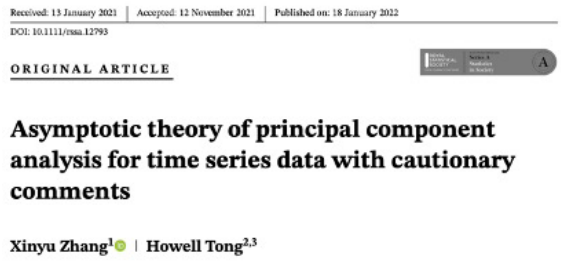


清华大学统计咨询中心受国家食品安全风险评估中心（以下简称 CFSA）周萍萍研究员委托，希望针对不同咖啡产品中呋喃（furan）及其衍生物的相关性分析问题给出具有优良统计学特性的解决方案。清华大学统计学研究中心邓柯副教授及周墨钦咨询师（2019级博士生）运用多元统计分析方法对不同咖啡产品中呋喃（furan）及其衍生物的潜在相关结构和异质模式进行了分析，协助食品安全专家更深入地认识了相关污染物在咖啡产品中的分布模式和规律。相关论文“Analysis of furan and its major furan derivatives in coffee products on the Chinese market using HS-GC–MS and the estimated exposure of the Chinese population”发表于食品科学技术领域的顶级期刊《Food Chemistry》（IF: 7.514；H-index: 221）。CFSA曹佩研究员为该文的第一作者，周萍萍研究员为通讯作者，邓柯副教授及周墨钦同学为共同作者参与了论文撰写。

本研究基于改进的顶空气相色谱 -质谱法（HS-GC-MS）分析了在中国市场上采集的咖啡样本中的呋喃及其衍生物的浓度，利用多元统计分析方法和可视化技术揭示了样本数据的内在结构，发现不同类型的咖啡产品的呋喃浓度水平和分布模式存在异质性，建议应加强对咖啡产品生产过程中的呋喃及其衍生物的控制研究。

中心博士生张心雨与汤家豪杰出访问教授在 JRSS-A发表时间序列数据主成分分析的研究论文

2022
05


Received: 13 January 2021 | Accepted: 12 November 2021 | Published on: 18 January 2022

DOI: 10.1111/rssa.12793

ORIGINAL ARTICLE

Asymptotic theory of principal component analysis for time series data with cautionary comments

Xinyu Zhang[1] | Howell Tong[2,3]

清华大学统计学研究中心 17级博士研究生张心雨与中心杰出访问教授汤家豪教授（Howell Tong）合作撰写的研究论文“Asymptotic theory of principal component analysis for time series data with cautionary comments”于今年年初正式发表于 Journal of the Royal Statistical Society: Series A (Statistics in Society)期刊。学术圈过往研究中经常直接对时间序列数据进行主成分分析而忽略其自相关性，针对这种误用，本文给出了时间序列主成分分析的统计推断性质和正确建模流程，并得出结论： 如果忽视数据间的相关性而直接进行统计推断，可能会对主成分的变量载荷做出误导性的过度解释。

主成分分析是统计学和数据科学中最常用的多元统计分析工具之一，但应用中也存在诸多误用现象。典型误用是： 对于时间序列数据，仍然使用独立数据假设下的理论结果。该论文强调了这种误用可能带来的问题。论文证明了时间序列主成分分析下的特征值和特征向量的中心极限定理，并给出其协方差的估计方法。论文关注方差比例和主成分载荷的统计推断，前者决定了主成分的数量，后者有助于主成分含义的解释。论文的研究结果表明： 在这种误用下，方差比例的统计推断仍然较为可靠，但是主成分载荷的统计推断会产生较大变化。论文着眼于一个投资组合管理的实例分析，以此提供了时间序列数据正确使用主成分分析的流程和案例。

- 专利及软著 -

【发明专利】

- 获批时间：2021 年 12 月
- 专利名称：基于电子病历信息的中文疾病名称智能标准化方法与系统
- **发 明 人：邓柯，李祺，刘军**

- 获批时间：2022 年 3 月
- 专利名称：基于贝叶斯联合建模优化算法的超材料设计方法及设备
- **发 明 人：邓柯，杨洋，**季春霖

- 获批时间：2022 年 3 月
- 专利名称：基于贝叶斯协同优化算法的超材料设计方法及相关设备
- **发 明 人：邓柯，杨洋，**季春霖

- 获批时间：2022 年 6 月
- 专利名称：一种双语句子自动对齐方法及装置
- **发 明 人：俞声，罗声旋**

【软件著作权】

盆底功能障碍性疾病智能诊断平台

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：邓柯课题组

登记日期：2021 年 7 月 30 日

OWAS 软件

著作权人：清华大学

研发团队：侯琳课题组

登记日期：2021 年 8 月 24 日

- 奖励荣誉 -

- 2020、2021 年度清华大学年度教学优秀奖
- 2021 北京高校第十二届青年教师教学基本功比赛二等奖（理科类 A 组）

- 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
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
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
北京市食品安全监控和风险评估中心	国家机关委托项目	2021 年 -2022 年	21 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
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 -2021 年	15 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	2020 年 -2023 年	120 万 (元)	邓　柯
上海起承文化发展有限公司	企业委托项目	2020 年 -2021 年	25 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
江苏先声医学诊断有限公司	合作研究	2020 年 -2022 年	50 万 （元）	侯　琳
清华大学万科公共卫生与健康学院	健康大数据支柱项目	2022 年 -2023 年	28 万 （元）	侯　琳

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 -2025 年	50 万 （元）	俞　声
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	72 万 （元）	俞　声
国家自然科学基金	青年项目	2019 年 -2021 年	24 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	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)
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)
国家自然科学基金	青年项目	2022 年 -2024 年	30 万 （元）	杨朋昆
国家自然科学基金	青年项目	2022 年 -2024 年	30 万 （元）	胡志睿

学术活动

- 主办学术活动 -

2021清华大学统计学与数据科学青年学者论坛

2021 年 10 月 23 日，清华大学统计学研究中心举办“2021 清华大学统计学与数据科学青年学者论坛”。 论坛旨在促进国内青年统计和数据科学学者间的学术交流与合作，更好地推动统计学和数据科学的发展，同时加强与兄弟院校间的协同合作。会议由清华大学统计学研究中心吴未迟副教授、杨朋昆助理教授及胡志睿助理教授发起并主持。来自清华大学、北京大学、中国人民大学、复旦大学、华东师范大学、浙江大学及上海财经大学等四十余名学者通过线上线下结合方式出席本次会议。

解放军总医院研究生院统计学与流行病学教研室 -清华统计合作研讨会

2022 年 2 月 18 日，解放军总医院研究生院统计学与流行病学教研室赛晓勇主任带领团队到访清华大学统计学研究中心，与中心生物健康统计团队师生座谈。双方在师资专家库共享、人才培养、科学研究等领域展开探讨，酝酿初步合作意向。

第六届北大 -清华统计论坛

2022 年 6 月 16 日，“第六届北大 - 清华统计论坛”成功举办。北大 - 清华统计论坛是北大、清华两校统计学科的传统学术活动，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，至今已成功举办六届。

随着两校统计学科的发展和人才队伍的壮大，北大 - 清华统计论坛的参会者逐年增加，本届论坛累计共有两百余人参会，除清北两校师生外还受到了很多其他高校和业界的学者关注。在特殊时期，两校统计学科的师生通过线上平台“云见面”及交流，活动精彩依旧。清华大学张学工教授和北京大学的丁剑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：清华大学统计学研究中心 17 级博士研究生李杰、北京大学统计科学中心 17 级博士研究生杨莹。

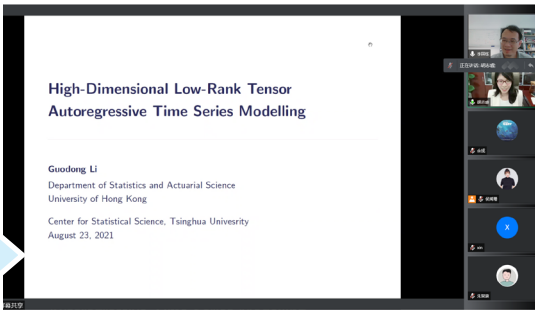
本年度“优秀海报奖”获得者： 清华大学统计学研究中心 18 级博士研究生朱珂、19 级博士研究生郑思捷、北京大学数学科学学院 18 级博士研究生王惠远、统计科学中心 18 级博士研究生邵凌轩。

- 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
2021.07.24	邬荣领	宾州州立大学统计遗传中心教授	Learning High-order Dynamical Interactome Networks from Big Static Data
2021.08.23	李国栋	香港大学统计精算系教授	High-Dimensional Low-Rank Tensor Autoregressive Time Series Modelling
2021.09.06	朱　柯	香港大学统计精算系助理教授	How Effective is the Regional Joint Environmental Policy in China? Evidence from Inverse Difference-in-Differences
2021.09.23	邱俊业	香港中文大学统计系副教授	Burn-in Selection in Simulating Time Series
2021.09.24	张　驰	中国科学院古脊椎动物与古人类研究所副研究员	贝叶斯全证据支端定年方法及应用
2021.10.11	杨建荣	中山大学中山医学院教授	Developmental Cell Lineage Trees, and the Quantitative Comparisons Between Them
2021.10.18	洪永淼	中国科学院大学经济与管理学院教授	Estimating and Testing for Smooth Structural Changes in Moment Condition Models
2021.11.01	谢尚宏	西南财经大学统计学院讲师	Integrative Network Learning for Multi-modality Biomarker Data
2021.11.08	马彦源	宾州州立大学统计系教授	Robust and Efficient Estimation under Nonignorable Missing Response
2021.11.15	Lucas Janson	哈佛大学统计系助理教授	Floodgate: Inference for Model-free Variable Importance
2021.11.22	龚若玢	罗格斯大学统计系助理教授	Towards Good Statistical Inference from Differentially Private Data
2021.11.29	Julia Palacios	斯坦福大学统计系助理教授	Distance-based Summaries and Modeling of Evolutionary Trees
2021.12.06	占　翔	北京大学生物统计系副教授	Statistical Methods for Microbiome Association Analysis
2021.12.13	洪　川	杜克大学生物统计与生物信息学系助理教授	Realizing the Potential of EHR Data for Clinical Research: Overcoming Noisiness, Privacy Constraints and Heterogeneity

时间	主讲人	工作单位及职位	报告题目
2021.12.14	黄东明	新加坡国立大学统计与数据科学系助理教授	Controlled Variable Selection with More Flexibility
2021.12.20	Martin Wainright	伯克利大学统计系教授	Beyond Worst-case: Instance-dependent Optimality in Reinforcement Learning
2022.02.28	周亚虹	上海财经大学经济学院教授	Nonparametric Identification and Estimation of the Generalized Additive Model
2022.03.07	周一鸣	角井 (北京） 生物技术有限公司创始人	人工智能技术（AI）在抗体药开发中的应用
2022.03.14	冯兴东	上海财经大学统计与管理学院教授	Quantile Regression for Nonignorable Missing Data with its Application of Analyzing Electronic Medical Records
2022.03.21	阮　丰	加州大学伯克利分校博士后	Designing Better Nonconvex Model for Modern Statistical Applications
2022.04.04	江　非	加州大学旧金山分校统计系助理教授	Time-varying Dynamic Network Model for Dynamic Resting State Functional Connectivity in fMRI and MEG Imaging
2022.04.11	吴　迪	北卡罗来纳大学教堂山分校生物统计系副教授	Novel Statistical Methods for Integrative Analysis of Metagenome, Metatranscriptome and Metabolome Applied in a Cohort of Early Childhood Caries (ECC)
2022.04.18	邹国华	首都师范大学数学科学学院教授	Asymptotic Distribution Theory for Model Averaging based on Information Criterion
2022.04.25	张景昭	清华大学交叉信息研究院助理教授	Convergence in Deep Learning does not Require Finding Stationary Points
2022.05.09	陈　豪	加州大学戴维斯分校统计系副教授	A Universal Nonparametric Event Detection Framework for Modern Data
2022.05.16	陆致用	美国国立卫生研究院高级研究员	PubMed & Beyond: Biomedical Text Mining for Knowledge Discovery
2022.05.23	赵子锋	圣母大学商学院助理教授	Optimal Change-point Testing for High-dimensional Linear Models with Temporal Dependence
2022.06.06	蒋斐宇	复旦大学管理学院青年副研究员	A Consistent Pivotal Specification Test
2022.06.13	黄　薇	墨尔本大学数学与统计学院讲师	Nonparametric Estimation of the Continuous Treatment Effect with Measurement Error Copy

香港大学统计与精算系
李国栋教授线上特邀报告

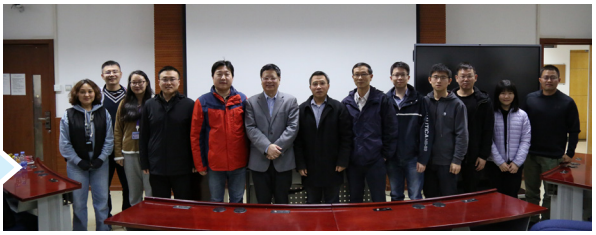


宾州州立大学统计遗传中心邬荣领
教授访问我中心，并做特邀报告

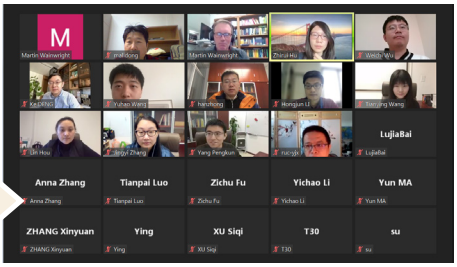
中国科学院古脊椎动物与古人类研究所张驰副研究员访问我中心并做学术报告

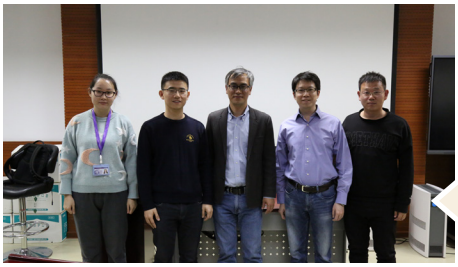
清华大学交叉信息研究院张景昭助理教授访问我中心，并做学术报告

中国科学院大学经济与管理学院洪永淼教授访问我中心，并做特邀报告

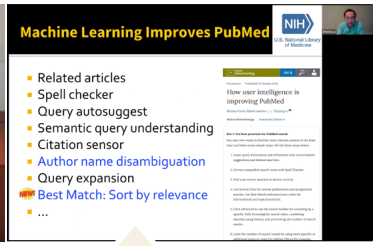


伯克利大学统计系
Martin Wainright教授
线上特邀报告





角井(北京)生物技术有限公司创始人周一鸣博士访问我中心，并做数据科学论坛报告



首都师范大学数学科学学院邹国华
教授线上特邀报告

美国国立卫生研究院高级研究员
陆致用博士线上特邀报告

- 参加学术活动 -

杨立坚

- 2022年06月　中国·线上　上海财经大学统计与管理学院 2022复杂函数型数据分析国际研讨会 (2022 International Workshop on Complex Functional Data Analysis)（邀请报告）
- 2021年12月　中国·线上　四川大学数学学院统计学科发展论坛
- 2021年11月　中国·线上　中国科学院数学与系统科学研究院第十三届系统科学发展论坛（大会报告）
- 2021年11月　中国·线上　上海财经大学统计与管理学院（讲座报告）
- 2021年10月　中国·线上　南京审计大学统计与数据科学学院（讲座报告）
- 2021年09月　美国·线上　国际泛华统计学会，ICSA Applied Statistics Symposium (邀请报告)
- 2021年07月　荷兰·线上　国际统计学会ISI，63-rd World Statistics Congress (邀请报告)

邓　柯

- 2022年05月　中国·深圳　南方科技大学统计与数据科学系（邀请报告）
- 2022年05月　中国·线上　北大人民医院（邀请报告）
- 2022年02月　中国·长春　长春工业大学数学与统计学院（邀请报告）
- 2021年11月　中国·北京　第二十一次全国统计科学讨论会（邀请报告）
- 2021年11月　中国·北京　中国科学院数学与系统科学研究院邹至庄讲座（邀请报告）

李　东

- 2022年06月　中国·线上　浙江大学数据科学中心（邀请报告）
- 2022年05月　中国·线上　南方科技大学统计与数据科学学术研讨会
- 2021年07月　中国·贵阳　第十二届全国概率极限理论和统计大样本理论学术研讨会

侯　琳

- 2022年04月　中国·北京　首都师范大学交叉科学研究院（邀请报告）
- 2022年03月　美国·线上　The University of Tennessee Health Science Center（邀请报告）
- 2022年01月　新加坡·线上　National University of Singapore（邀请报告）
- 2021年11月　中国·线上　河南大学（邀请报告）
- 2021年11月　中国·北京　中国科学院数学与系统数学研究院（邀请报告）
- 2021年10月　中国·昆明　中国数学会2021年学术年会（分组报告）

俞　声

- 2021年 12月　中国·线上　CHIP2021 第七届中国健康信息处理大会（大会报告）
- 2021年 11月　中国·深圳　IDEA大会（发言）

刘汉中

- 2021年12月　中国·线上　智源因果社区（邀请报告）
- 2021年11月　中国·北京　中国人民大学（邀请报告）
- 2021年09月　中国·北京　北京大学生物统计系（邀请报告）

林　乾

- 2021年12月　中国·北京　中国人民大学统计学院（邀请报告）
- 2021年11月　中国·线上　中国社会科学院大学青年学者论坛（邀请报告）

吴未迟

- 2022年06月　中国·线上　浙江大学数据科学中心（邀请报告）
- 2022年06月　日本·线上　EcoSta 2022 – CMStatistics（组织session+邀请报告）
- 2021年12月　英国·线上　CFE-CMStatistics 2021
- 2021年11月　中国·线上　吉林大学数学系（邀请报告）

张静怡

- 2022年04月　美国·线上　亚利桑那大学流行病与生物统计系(邀请报告)
- 2021年09月　中国·线上　武汉大学数学与统计学院（邀请报告）
- 2021年08月　美国·线上　2021 Joint Statistical Meetings - Medical Devices and Diagnostics Section (大会报告)
- 2021年07月　中国·成都　第三届全国大数据与人工智能科学大会 (大会报告)
- 2021年07月　中国·长沙　湖南大学信息科学与工程学院（邀请报告）
- 2021年07月　中国·长沙　湖南大学数学学院（邀请报告）
- 2021年07月　中国·长沙　国防科技大学文理学院系统科学系（邀请报告）
- 2021年07月　中国·北京　狗熊会学术报告（邀请报告）

杨朋昆

- 2022年04月　中国·线上　南方科技大学统计与数据科学系（邀请报告）
- 2021年10月　中国·北京　2021清华大学统计学与数据科学青年学者论坛（组织者）
- 2021年10月　中国·北京　百度研究院大数据实验室（邀请报告）
- 2021年10月　中国·线上　香港大学统计与精算学系 （邀请报告）
- 2021年09月　中国·北京　中国科学院数学与系统科学研究院 （邀请报告）
- 2021年08月　美国·线上　Conference on Learning Theory （发言）

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》

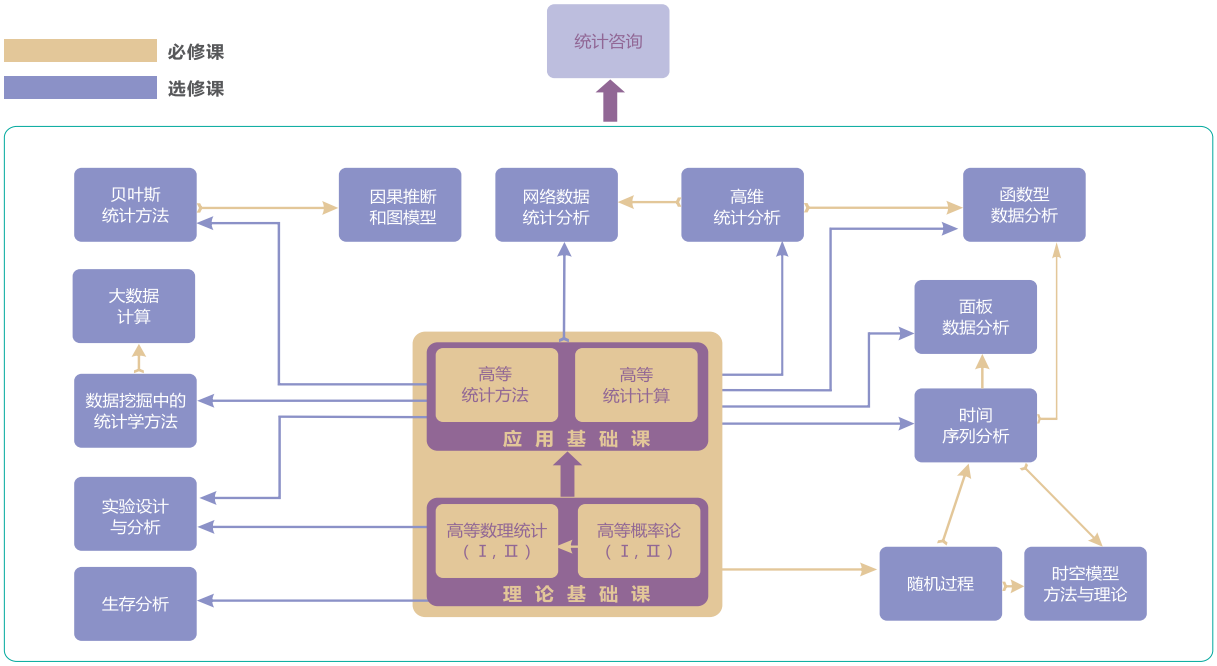
2018 年，清华大学统计学研究中心对标国际顶尖院校的统计学通识课程，筹备建设了清华大学首门统计学通识课，以应对大数据时代各专业领域对统计学的强烈需求。课程不但面向全校各专业本科学生开放，更加入清华 - 北大互选课名单，预留部分名额，向兄弟院校的学生开放。课程由中心长聘副教授邓柯、李东两位老师讲授。内容包括统计学基本思想、基础理论和基本方法的理论，横跨工程、生物医学、经济金融、人文社科等多个专题，融合数据分析实践和统计编程入门的实践。

- 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有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。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。

课程设置





博士生论坛

2021 年 11 月 6 日，由清华大学统计学研究中心发起并组织的“2021 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
	第七届全国高校研究生统计论坛“十佳论文”奖	钟　晨	Inference and Prediction for ARCH Time Series via Innovation Distribution Function	2021 年 11 月	杨立坚
	2021年清华大学工业工程系“未来教授培养计划”	宋泽宁	/	2021 年 12 月	杨立坚
	2021 年清华大学工业工程系“未来教授培养计划”	李　杰	复杂时间序列的统计推断理论及预测方法	2021 年 12 月	杨立坚
	2022 年清华大学优秀博士学位论文	李　杰	复杂时间序列的统计推断理论及预测方法	2022 年 6 月	杨立坚
	2022 年清华大学毕业生启航奖铜奖	李　杰	复杂时间序列的统计推断理论及预测方法	2022 年 6 月	杨立坚
	第六届北大 - 清华统计学论坛优秀毕业生	李　杰	复杂时间序列的统计推断理论及预测方法	2022 年 6 月	杨立坚
	第六届北大 - 清华统计论坛优秀海报奖	郑思捷	Inference for Dependent Error	2022年 6月	杨立坚

	奖项名称	获奖者	具体信息	时间	导师
	2021 百济神州青年优秀论文二等奖	韩庭萱	Detection of Blood Stains Using Computer-vision based Algorithms and Its Association with Postoperative Outcomes in Thoracoscopic Lobectomies	2021 年 12 月	邓　柯
	2021 年清华之友 - 东丽英才奖学金	陶宇心	/	2021 年 10 月	李　东
	2021 年清华大学工业工程系“未来教授培养计划”	张心雨	/	2021 年 12 月	李　东
	2021 百济神州青年优秀论文	宋　爽	A Data-adaptive Bayesian Regression Approach for Accurate Polygenic Risk Prediction	2021 年 12 月	刘　军 侯　琳
	2021 年清华大学工业工程系“未来教授培养计划”	宋　爽	A Data-adaptive Bayesian Regression Approach for Accurate Polygenic Risk Prediction	2021 年 12 月	刘　军 侯　琳
	2021年国家奖学金	宋　爽	A Data-adaptive Bayesian Regression Approach for Accurate Polygenic Risk Prediction	2021年 12月	刘　军 侯　琳
	2021年清华大学蒋南翔奖学金	宋　爽	A Data-adaptive Bayesian Regression Approach for Accurate Polygenic Risk Prediction	2021年 12月	刘　军 侯　琳

	奖项名称	获奖者	具体信息	时间	导师
	2021年清华大学一二·九辅导员郭明秋奖	宋　爽	A Data-adaptive Bayesian Regression Approach for Accurate Polygenic Risk Prediction	2021年12月	刘　军 侯　琳
	第十届全国生物信息学与系统生物学学术大会墙报二等奖	余　博	Differential Analysis of RNA Structure Probing Experiments at Nucleotide Resolution: Uncovering Regulatory Functions of RNA Structure	2021年9月	侯　琳
	2021百济神州青年优秀论文	余　博	Differential Analysis of RNA Structure Probing Experiments at Nucleotide Resolution: Uncovering Regulatory Functions of RNA Structure	2021年12月	侯　琳
	2021清华之友-丰田奖学金	朱　珂	Pair-switching Rerandomization	2021年10月	刘汉中
	2021百济神州青年优秀论文	朱　珂	Pair-switching Rerandomization	2021年12月	刘汉中
	2021年清华大学工业工程系“未来教授培养计划”	朱　珂	Pair-switching Rerandomization	2021年12月	刘汉中

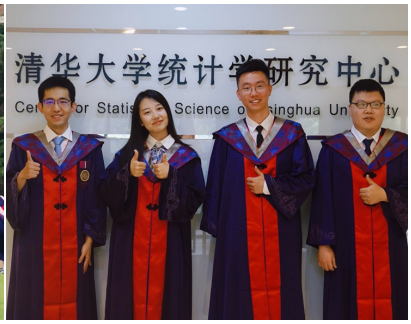
	奖项名称	获奖者	具体信息	时间	导师
	第六届北大-清华统计论坛优秀海报奖	朱　珂	Blocking, Rerandomization, and Regression Adjustment in Randomized Experiments with High-dimensional Covariates	2022年6月	刘汉中
	2021清华之友-东丽英才奖学金	白露佳	Testing for Long-range Dependence in Non-stationary Time Series Time-varying Regression	2021年10月	吴未迟
	2021统计理论及应用国际研讨会优秀论文	白露佳	Testing for Long-range Dependence in Non-statiionary Time Series Time-varying Regression	2021年12月	吴未迟

毕业生专栏

统计学研究中心 2021 年 7 月 -2022 年 6 月共 7 名同学顺利毕业：

2022 届　徐嘉泽、单娜阳、郭瀚民、李杰、袁正、张心雨、钟晨





徐嘉泽　　导师：邓柯 / 毕业去向：某私募基金公司

读博五年多的时间，有幸见证了统计中心发展壮大的过程。感谢中心提供的平台，雄厚的师资力量、宽松又严谨的学术氛围、一流的软硬件设施，可以让博士生在纷杂的社会现状下，静下心来研究学术。读博是一个磨炼心志的过程，不断试错、保持恒心，在人类已有的认知边界上，不断探索；读博的生活也是一种享受，享受付出的过程、时刻思考的精神状态、耐得住寂寞的坚持和探索精神，还可以享受收获的果实。祝愿中心的发展越来越好，感谢中心老师和同学对我的帮助和关心，感谢导师对我的指导和培养！

单娜阳　　导师：侯琳 / 毕业去向：首都经济贸易大学讲师

很感谢有机会在清华统计学研究中心度过了难忘的岁月。感谢统计中心为同学们提供了顶尖的师资力量、丰富的学术资源和自由的科研氛围。统计中心拥有一支专业卓越的研究团队，一支精干高效的行政团队，还有一群朝气蓬勃的博士生，感谢遇见你们。在此特别感谢我的导师侯琳老师对我科研的指导以及生活中给予的温暖。侯老师治学严谨、谦逊平和，她的言传身教将使我受益终生！最后祝愿统计中心成为国内统计学科的中心，国际统计学界的重镇。祝愿老师们、同学们前程似锦。

郭瀚民　　导师：侯琳 / 毕业去向：斯坦福大学博士后

毕业在即，回顾自己在统计中心五年的直博生活，感觉获益颇多。在五年的时间里，中心给各位同学提供了一个自由、开放、包容的科研环境，同学们可以尽情发挥自己的学术才华。每周的统计学论坛邀请海内外知名学者做报告也极大地拓宽了我们的学术视野。中心的老师和同学们一起奋斗在科研的第一线，同学们耳濡目染，对学术的兴趣又浓了一分，科研的能力也持续进步。衷心希望统计中心越办越好，在未来持续培养出统计学领域的优秀人才。

李　杰　　导师：杨立坚 / 毕业去向：中国人民大学师资博士后

经常感慨时光已逝，转眼间我已经博士毕业，结束了在清华园的五年学习生活。我非常感谢我的导师杨立坚老师。我一直觉得很幸运能够跟着杨老师读博，杨老师扎实的数理功底、严谨的科研态度、富有创造力的学术思维令我敬佩。平日里他对我悉心指导，关怀鼓励，他春风化雨般的言传身教让我在学术道路上逐渐步伐坚定，眼界开阔，信心满满。杨老师不仅是我的学术导师，更是我人生的引路人。这五年我见证了统计中心的规模壮大，从 10 余个学生发展到 70 多名师生的大集体，也亲历了统计中心日新月异般的发展进步。感谢统计中心为我们提供了良好的科研平台、丰富的学术资源和前沿的教学课程，拓宽了我们的学术视野。

我希望自己始终坚持“面向应用，背靠理论，写好算法”的统计学思想，不跟风，不浮躁，做理论扎实、应用价值突出的统计学研究。我时刻谨记自己的责任和使命，努力践行“自强不息，厚德载物”的校训，为成为一名合格的高校教师，一名于国家、于民族有用的清华人而努力。

袁　正　　导师：俞声 / 毕业去向：阿里巴巴达摩院

在统计中心的五年如白驹过隙。现在我还能回忆起参加统计中心夏令营时的场景，当时的我们还非常稚嫩，对如何做科研充满兴趣却也不甚了解。经过统计中心这五年的培育，我们逐渐学会统计理论，学会如何开展科研，学会撰写论文;是统计中心锻炼我们成为了独立的研究者。在最后，祝统计中心越做越好，祝各位老师和同学发展顺利，学业有成!

张心雨　　导师：李东 / 毕业去向：爱荷华大学博士后

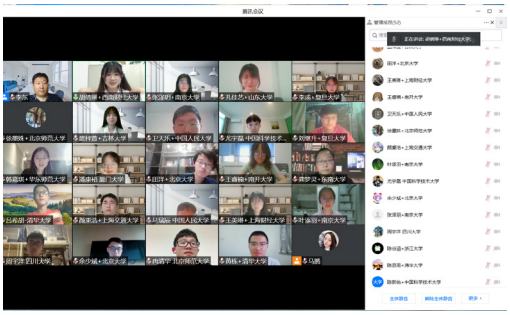
时间匆匆，即将结束在清华的五年求学。五年来经历了很多：刚入学时的孤独、不适应，新冠疫情对生活的巨大影响，毕业前的焦虑和抉择，等等。但也收获了更多:能感知到自己在学术上逐渐进步，结识了太多令我敬佩的人，得到了太多不计回报的帮助。生活上，我特别想感谢学校，每当我使用校内便捷的医疗、运动、艺术、图书资源时，都会倍感珍惜。科研上，我特别想感谢统计中心，中心提供了前沿的学术资源和充满活力的科研环境。我尤其想感谢我的导师李东老师，李老师不仅手把手指导我的科研，也尽心帮我解决生活中的困难，耐心疏导我、鼓励我。我想，这是我的幸运，能得到学校和中心的栽培，能得到渊博又宽厚的老师的指导，能结识一群努力纯粹又优秀的同学。未来的日子，衷心祝愿统计中心越办越好。

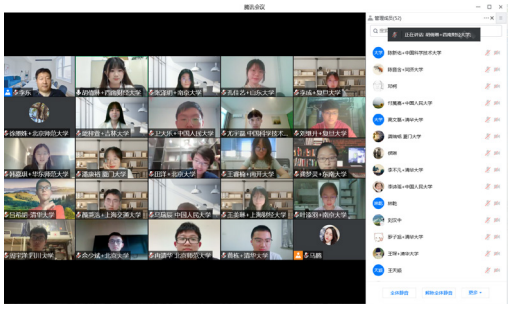
钟　晨　　导师：杨立坚 / 毕业去向：武汉大学博士后

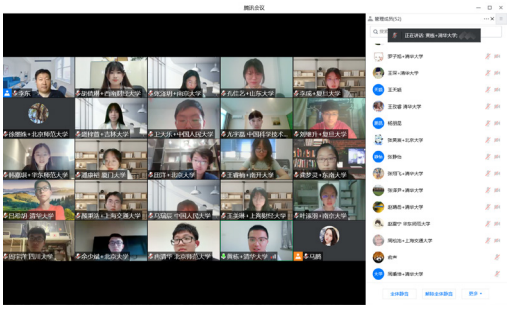
清华大学统计学研究中心各位老师们卓越的教学、研究和管理工作，为我们营造了良好的科研氛围。在中心的学习阶段，各类区域性、国际性的学术会议以及每周的统计学论坛，不仅极大地拓宽了我们的视野，让我们能够了解前沿的学术动态，也为我们提供了宽广的交流平台。导师杨立坚老师的悉心教导，让我受益匪浅，我将铭记于心。离开中心，步入人生新的阶段，今后我将恪守“自强不息，厚德载物”的校训，做一个踏实严谨、积极探索信念坚定的科研工作者。我也希望通过自己的身体力行将中心的精神传承延续。聚散终有时，源头活水来。在此也恭祝中心能够越办越好，积聚与培育出更多的优秀人才。

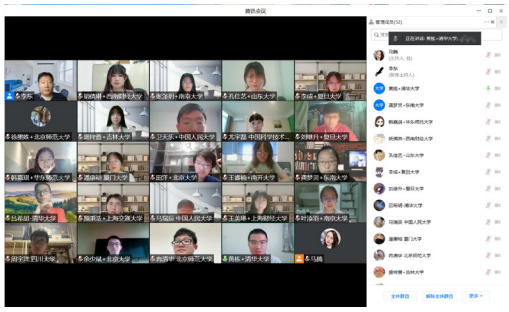
-2022 清华大学统计学研究中心优秀大学生夏令营 -

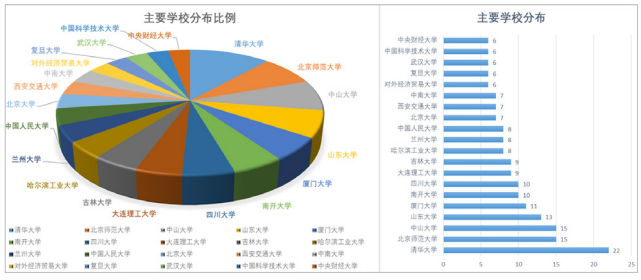
2022 年 5 月 26-29 日，清华大学统计学研究中心优秀大学生夏令营活动成功举办。本届夏令营共收到来自 95 所高校的 351 名同学的申请，经过层层筛选，来自清华大学、北京大学、北京师范大学、中国人民大学、中国科学技术大学、吉林大学、上海交通大学等多所国内知名高校的共 42 名同学成功入营。











社会服务及影响

- 社会服务 -

杨立坚

- 2022 年 05 月　中国国家自然科学基金评审专家
- 2021 年 12 月　北京理工大学新进教师同行评议专家
- 2021 年 11 月　复旦大学博导评审专家
- 2021 年 10 月　中国统计学会第一届统计科学技术进步奖评审专家
- 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)
- 2020 年 07 月－至今　清华大学理科学术委员会委员
- 2020 年 07 月－至今　科学出版社统计与数据科学系列丛书编委
- 2020 年 07 月－至今　苏州市现场统计研究会第六届理监事会顾问
- 2020 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
- 2020 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
- 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
- 2018 年 11 月－ 2021 年 11 月　西安交通大学全球健康研究院兼职教授

邓　柯

- 2022 年 06 月　教育部“海外优青” 通讯评审评委
- 2022 年 04 月－ 2025 年 03 月　清华大学质量与可靠性研究院副院长
- 2021 年 12 月　之江实验室开放课题评审专家
- 2021 年 12 月　教育部第五轮学科评估评估专家
- 2021 年 12 月　丘成桐中学生数学竞赛评委
- 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 09 月－ 2024 年 09 月　清华大学求真书院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　中国国家自然科学基金评审专家
- 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03月－至今　中国现场统计研究会计算统计分会常务理事、秘书长

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
  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

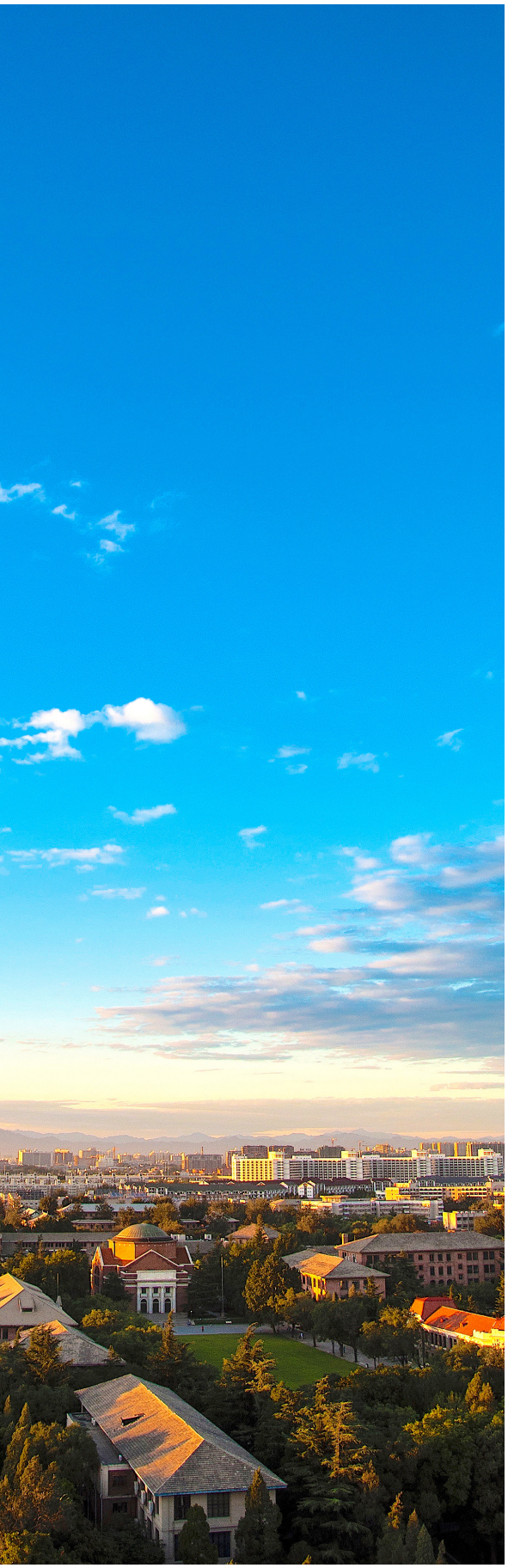
- 2019年－至今　全国工业统计学教学研究会青年统计学家协会第一届理事会理事
- 2019年－至今　北京应用统计学会理事
- 2016年－至今　中国现场统计研究会计算统计分会副秘书长

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

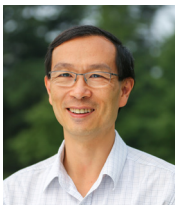
吴未迟

- 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



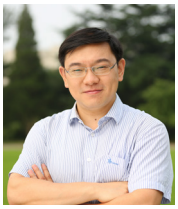
- 学术杂志服务 -

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。



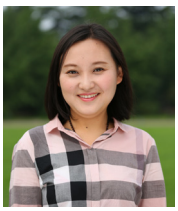
杨立坚　教授

- *Stat* 副主编
- *Statistica Sinica* 副主编
- *Brain Sciences* 编　委



邓　柯　副教授

- *Statistica Sinica* 副主编
- *Series in Biostatistics* 编　委
- 《统计与精算》 编　委
- 《数字人文》 编　委
- 《应用概率统计》 编　委
- 《应用数学和力学》 编　委



侯　琳　副教授

- *Statistics in Biosciences* 副主编
- *Quantitative Biology* 编　委

统计咨询中心概况

清华大学统计咨询中心由清华大学统计学研究中心发起并成立，于 2017 年 5 月开始运行，面向清华各院系及社会各界提供数据分析和统计咨询服务。

统计咨询中心一方面为清华大学各院系的广大师生、校外实体及企业提供高水平的统计建议与支持；另一方面，以跨专业为基础，通过合作鼓励统计学者和其他领域人员之间的交叉研究； 同时，培养统计学研究生与其他领域研究者的交流互动能力，成为能够运用统计学技巧解决实际问题的高效合作者。

业务团队

邓　柯　主任

清华大学统计学研究中心长聘副教授

高级咨询师　科研助理　学生咨询师

高梦昭 博士

王江典 博士

宋希婷 硕士

任吉杨

清华大学统计学研究中心博士研究生

王海沣

清华大学统计学研究中心博士研究生

周墨钦

清华大学统计学研究中心博士研究生

学术顾问

Donald Rubin 教授

- 清华大学特聘教授
- 美国科学院院士

Wing Wong 教授

- 斯坦福大学统计系终身教授
- 美国科学院院士

许宪春 教授

- 清华大学经济管理学院教授
- 原国家统计局副局长

访问学者

牛晓月 副教授

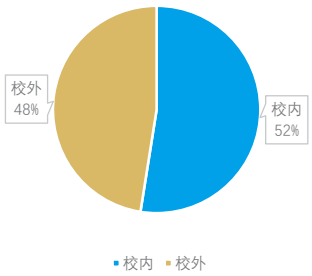
美国宾夕法尼亚州立大学统计系副教授

年度业务总览

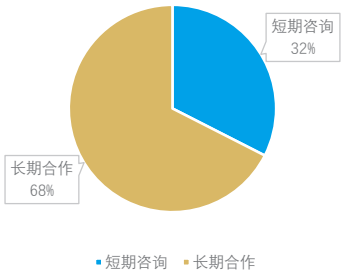
本年度统计咨询中心共承接咨询案例 40 例，包括秋季学期 21 例，春季学期 19 例。其中校内咨询 21 例，占比 52%；校外咨询 19 例，占比 48%。在校内咨询中，短期咨询占比 38.1%，长期合作占比 61.9%；校外咨询中短期咨询占比 26.3%，长期合作占比 73.7%。无论项目源于校内还是校外，长期合作的占比高于短期咨询，说明咨询中心的业务已经转向了更加深入密切的合作方式，对我们的专业和业务能力都提出了更高要求。

咨询中心持续为校内外客户提供高质量的专业咨询服务。客户组成也更加多元化，特别是校外咨询，本年度客户来自政府、高校、医院、科研单位、公司等各行各业。

2021~2022学年项目来源



2021~2022学年项目类型



校内项目来源	项目数量
车辆学院	1
环境学院	1
建筑学院	3
教评中心	5
教务处	1
精密仪器系	1
精仪系	1

校内项目来源	项目数量
软件学院	1
社科学院	1
数学系	1
体育部	1
写作中心	1
信息办	2
长庚医院	1

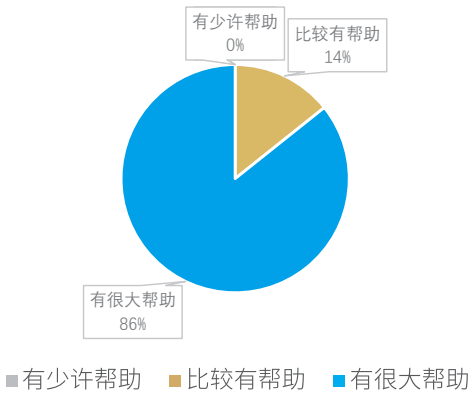
校外项目来源	项目数量
高校	4
公司	4
科研单位	3
医院	2
政府	6

满意度调查

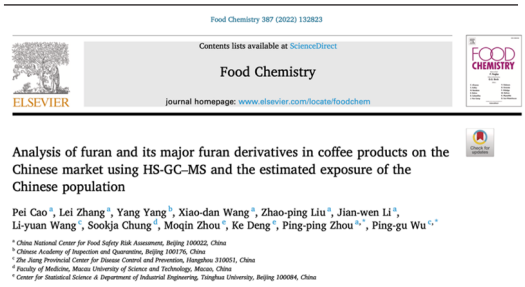
为了提高咨询服务质量，完善服务内容，在咨询结题后对客户进行服务满意度调查。根据本学年调查反馈，咨询客户对咨询中心服务满意度高达 100%，其中 86% 的客户认为咨询中心提供的咨询建议有很大帮助。

	咨询整体评价	咨询会议评价	会议纪要评价
很满意	100%	100%	100%
满意			
一般			
不满意			
很不满意			

满意度调查-咨询建议评价



发表论文


Food Chemistry 387 (2022) 132823

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Analysis of furan and its major furan derivatives in coffee products on the Chinese market using HS-GC–MS and the estimated exposure of the Chinese population

Pei Cao [a], Lei Zhang [a], Yang Yang [b], Xiao-dan Wang [a], Zhao-ping Liu [a], Jian-wen Li [a], Li-yuan Wang [c], Sookja Chung [d], Moqin Zhou [e], Ke Deng [c], Ping-ping Zhou [a,*], Ping-gu Wu [c,*]

[a] China National Center for Food Safety Risk Assessment, Beijing 100022, China
[b] Chinese Academy of Inspection and Quarantine, Beijing 100176, China
[c] Zhe Jiang Provincial Center for Disease Control and Prevention, Hangzhou 310051, China
[d] Faculty of Medicine, Macau University of Science and Technology, Macao, China
[e] Center for Statistical Science & Department of Industrial Engineering, Tsinghua University, Beijing 100084, China


Journal of Environmental Psychology
Volume 79, February 2022, 101745

Using natural intervention to promote subjective well-being of essential workers during public-health crises: A Study during COVID-19 pandemic

Chenhao Hu [a], Ke Zhu [b], Kun Huang [b], Bo Yu [b], Wenchen Jiang [c], Kaiping Peng [a], Fei Wang [a, d, *]


Neuroscience Informatics
Volume 2, Issue 2, June 2022, 100047

Original article

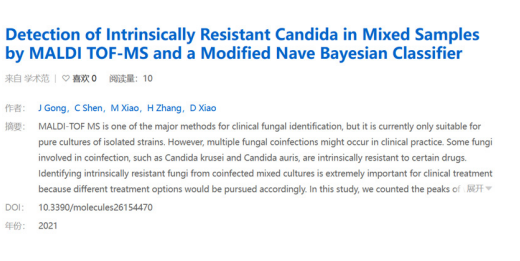
The use of a new classification in endovascular treatment of dural arteriovenous fistulas

Huachen Zhang [a], Ke Zhu [b], Jiangdian Wang [b], Xianli Lv [a]


线上和线下本科教学质量的比较分析——基于清华大学教学评估数据

Quantitative Analysis and Undergraduate Teaching Quality Comparison Between Online Teaching and In-class Teaching——Based on Teaching Evaluation Data from Tsinghua University


Bootstrap多重插补在填补医学研究缺失数据中的应用


Detection of Intrinsically Resistant Candida in Mixed Samples by MALDI TOF-MS and a Modified Nave Bayesian Classifier

来自 学术范 | 喜欢 0 阅读量：10

作者： J Gong, C Shen, M Xiao, H Zhang, D Xiao

摘要： MALDI-TOF MS is one of the major methods for clinical fungal identification, but it is currently only suitable for pure cultures of isolated strains. However, multiple fungal coinfections might occur in clinical practice. Some fungi involved in coinfection, such as Candida krusei and Candida auris, are intrinsically resistant to certain drugs. Identifying intrinsically resistant fungi from coinfected mixed cultures is extremely important for clinical treatment because different treatment options would be pursued accordingly. In this study, we counted the peaks of 展开

DOI： 10.3390/molecules26154470

年份： 2021

共同发表 1:

Pei Cao, Lei Zhang, Yang Yang, Xiao-dan Wang, Zhao-ping Liu, Jian-wen Li, Li-yuan Wang, Sookja Chung, Moqin Zhou, Ke Deng, Ping-ping Zhou, Ping-gu Wu.(2022）. Analysis of furan and its major furan derivatives in coffee products on the Chinese market using HS-GC–MS and the estimated exposure of the Chinese population, Food Chemistry, Volume 387, 1 September 2022, 132823.

共同发表 2:

Hu, C., Zhu, K., Huang, K., Yu, B., Jiang, W., Peng, K., & Wang, F. (2022). Using natural intervention to promote subjective well-being of essential workers during public-health crises: A Study during COVID-19 pandemic. Journal of Environmental Psychology, 79, 101745.

共同发表 3:

Zhang, H., Zhu, K., Wang, J., & Lv, X. (2022). The use of a new classification in endovascular treatment of dural arteriovenous fistulas. Neuroscience Informatics, 2, 100047.

共同发表 4:

王江典，沈翀，杨蕾，高梦昭，王红，邓柯(2022)，线上和线下本科教学质量的比较分析——基于清华大学教学评估数据,《中国电化教育》,2022年第 3期 90-95,102,共 7页 .

共同发表 5:

裴敏玥 ,沈翀 ,李楠 & 赵一鸣 .(2022).Bootstrap 多重插补在填补医学研究缺失数据中的应用 . 中华儿科杂志 (01),2-2.

共同发表 6:

Gong, J., Shen, C., Xiao, M., Zhang, H., Zhao, F., Zhang, J., & Xiao, D. (2021). Detection of Intrinsically Resistant Candida in Mixed Samples by MALDI TOF-MS and a Modified Naïve Bayesian Classifier. Molecules (Basel, Switzerland), 26(15), 4470. https://doi.org/10.3390/molecules26154470

年度优秀咨询师

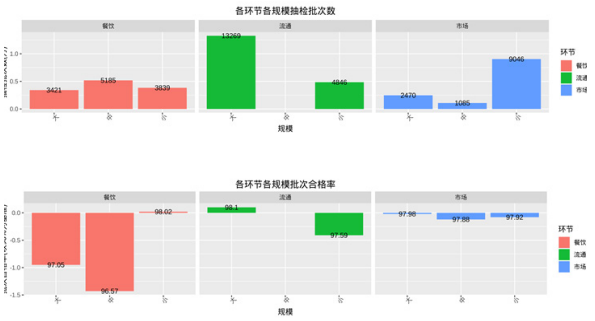
任吉杨　清华大学统计学研究中心博士研究生（2021.9-2022.7）

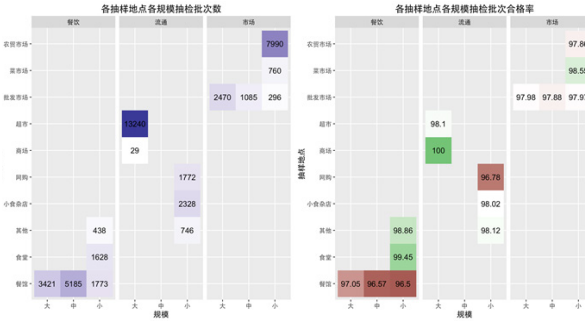
王海沣　清华大学统计学研究中心博士研究生（2021.9-2022.7）

典型案例

案例一：北京市食品安全抽检合格率统计方法研究

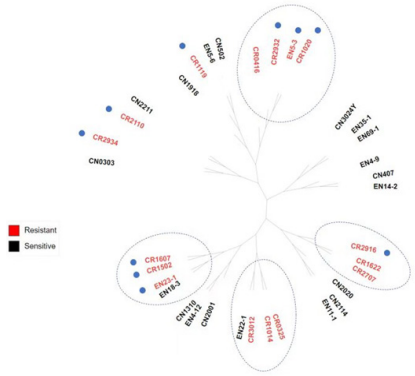
为了更加全面科学地评估市场食品安全状况，国家市场监督管理总局自 2017 年起不断推进食品安全评价性抽检工作，各地也陆续开展相应工作。为了更好地利用抽检数据，更加全面、客观反映北京市食品安全状况，北京市市场监督管理局委托清华大学统计学研究中心邓柯副教授统计咨询中心团队开展“北京市食品安全抽检合格率统计方法研究”，从多角度深入分析了北京市 2020 年度、2021 年度北京市食品安全情况，构建了食品安全评价理论模型，对北京市各区各食品种类的食品安全进行了有效评估。此外，团队还以准确评估食品安全评价指数为目标，设计了北京市 2022 年食品安全评价性抽检方案。研究成果帮助了相关部门全面掌握北京市食品安全情况，提升了监管效率，节约了政府资源。



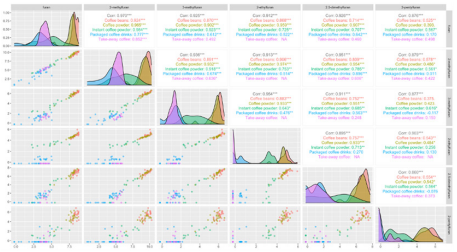


案例二 :真菌耐药性菌株分析

基于测序的耐药性发现对于治疗耐药菌感染以及限制耐药菌扩散的临床和公共卫生实践至关重要，但在技术与方法上仍面临一定的挑战。中科院微生物所吴琦博士团队联合清华大学统计学研究中心邓柯副教授统计咨询中心团队开展了“真菌耐药性菌株分析”研究，对耐药菌及同属敏感菌的 k-mer 进行分析，在不破坏遗传结构的基础上设计了带有限制的置换检验，以从中发现耐药基因。本研究提出的置换检验方法充分利用了 k-mer 频数的信息，并尽可能地保证了真菌遗传结构的不变性，对统计学方法在微生物领域的应用做出了一定的创新，推进了真菌耐药基因发现的研究。



案例三 :统计学在咖啡中污染物浓度相关结构分析中的应用



清华大学统计咨询中心受国家食品安全风险评估中心周萍萍研究员委托，希望针对不同咖啡产品中呋喃（furan）及其衍生物的相关性分析问题给出具有优良统计学特性的解决方案。清华大学统计学研究中心邓柯副教授及周墨钦咨询师利用多元统计分析方法对不同咖啡产品中呋喃（furan）及其衍生物的潜在相关结构和异质模式进行了分析，协助食品安全专家更深入地认识了相关污染物在咖啡产品中的分布模式和规律。本研究基于改进的顶空气相色谱 - 质谱法（HS-GC-MS）分析了在中国市场上采集的咖啡样本中的呋喃及其衍生物的浓度，利用多元统计分析方法和可视化技术揭示了样本数据的内在结构，发现不同类型的咖啡产品的呋喃浓度水平和分布模式存在异质性，建议应加强对咖啡产品生产过程中的呋喃及其衍生物的控制研究。

案例四 : 运用统计分析提升同位素核半径估计精度

关于同位素核半径的估计一直是学者广泛关注的问题，对同位素核半径的精确测量可以为核物理的理论推导提供有力的实验依据。清华精密仪器系联同清华大学统计学研究中心邓柯副教授统计咨询中心团队，合作开展了“同位素核半径估计”的项目工作，并运用统计分析的框架整合了现有的同位素的光谱测量数据，并利用测量误差模型（Measurement Error Model）提高了现有核半径估计的精度。课题组运用统计分析方法整合了现有多光谱多同位素下的测量数据，并利用文献中的相关结果，提高了同位素核半径估计的精度。通过极大似然估计与数值优化或者 EM 算法结合，得到多光谱多同位素下模型参数的估计值，进而实现已知光谱测量值下同位素核半径的推断。相关模拟实验也验证了该模型对同位素核半径估计的有效性。

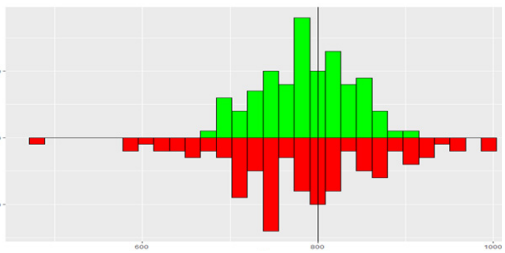
i: 同位素 \ j: 光谱	$j=1$	$j=2$	$j=3$
$i=1$	Y_{11}^{obs}	Y_{12}^{obs}	Y_{13}^{obs}
$i=2$	Y_{21}^{obs}	Y_{22}^{obs}	Y_{23}^{obs}
$i=3$	Y_{31}^{obs}	Y_{32}^{obs}	Y_{33}^{obs}
$i=4$	Y_{41}^{obs}	Y_{42}^{obs}	Y_{43}^{obs}
$i=5$(新数据)	Y_{51}^{obs}	Y_{52}^{obs}	Y_{53}^{obs}
参数	k_1, b_1	k_2, b_2	k_3, b_3

联合推断

t: 核半径
t_1^{obs}
t_2^{obs}
t_3^{obs}
t_4^{obs}
t_5=?

测量误差	$j=1$	$j=2$	$j=3$	t
$i=1$	σ_{11}	σ_{12}	σ_{13}	τ_1
$i=2$	σ_{21}	σ_{22}	σ_{23}	τ_2
$i=3$	σ_{31}	σ_{32}	σ_{33}	τ_3
$i=4$	σ_{41}	σ_{42}	σ_{43}	τ_4
$i=5$	σ_{51}	σ_{52}	σ_{53}	

原始数据　输入　输出



编　　辑：邓　柯
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

扫码关注微信公众号
水木数据派