



北京大学统计科学中心

Center for Statistical Science, Peking University

2017年度报告



2010 年 7 月统计中心第一次国际顾问和科学委员会会议

目 录

一、综述	2
二、中心要闻	3
三、人员组成	14
四、科研项目	22
五、博士毕业生去向	24
六、交流访问	25
七、论文发表	32
八、研究报告	36
九、学术讲座	39
十、委员会	45

北京大学统计科学中心（Center for Statistical Science, Peking University）成立于 2010 年 7 月，由数学科学学院、光华管理学院及其它从事统计研究的相关教师构成，是北京大学实行特殊机制的跨学院交叉学科研究机构。

统计科学中心依托北京大学深厚的学术底蕴和多学科的优势，整合全校统计学研究力量，致力于培养综合型的统计和数据科学人才，扩展统计学在其它学科中的应用，推动统计学研究与教育走向世界，建设国内领先、世界一流的研究机构。

统计科学中心的主要任务是建立一支具有国际水平的统计学研究教学团队，促进统计学与其他学科的交叉合作以及加强统计学在各学科中的应用，推动和提高各学科的实证与统计分析水平，为这些学科的发展提供数据收集和数据分析的统计方法。

中心设国际顾问委员会、科学委员会、执行委员会和指导委员会。国际顾问委员会负责提出有关中心发展方向等方面的建议，评估中心的工作和成果；科学委员会在人才引进和学术研究等方面给予意见和建议；执行委员会参与中心的管理运行工作。中心在这几年中利用新体制共引进海外高层次人才 7 人，其中 3 人入选中组部“国家级重大人才青年计划”，三人入选国家“国家级重大人才计划”长期项目，三人入选国家“国家级重大人才计划”短期项目。

统计中心的成立已在北京大学形成一个吸引海内外优秀统计人才的基地。统计中心将继续营造一个开放式、国际化的学术环境，广泛联系全球各地的统计学者，推动北大统计学“一流学科”建设，我们希望在 2030 年将北京大学的统计学科建设成为世界一流学科。



Peter G. Hall 教授摄

一、综述

2017 年统计科学中心继续在统计学研究、教员引进、博士生培养以及推动统计学发展等方面开展工作。

由统计中心牵头的北大统计学科入选教育部“世界一流建设”学科系列,全国有八个学校的统计学科入选。建设一流学科也成为统计中心下一个时期的主要任务。在 12 月 28 日教育部公布的全国第四轮学科评估结果中,北京大学统计学一级学科与中国人民大学共同获评 A+, 这充分体现了统计中心过去四年在整合全校统计研究力量方面所做的努力和学校对统计学科投入的回报。统计中心和光华管理学院合聘了苗旺博士,他将于 2018 年 8 月入职,这将是中心时隔四年后再次招聘教员。初秋,中心开始搬离了燕东园 32 号小楼,迁入理科五号楼。

2017 年统计中心积极开展同国际和国内统计学界的学术交流与合作。5 月 14 日至 16 日,第二届“大数据时代下的高维统计建模与分析研讨会”成功举行,本次会议由统计中心和数量经济与数理金融教育部重点实验室共同主办,北京大学光华管理学院承办,来自全国各高校近二百名师生参加了会议。5 月底统计中心成功举办了第三届优秀大学生夏令营,十几所高校的 25 名三年级本科生参加,我们从中录取了四名博士研究生。6 月 13 日至 14 日,第十三届计量经济学理论与应用国际研讨会成功举办,本次会议吸引了来自世界各地的优秀学术文章,该会议现已成为亚太地区计量经济学界的领先学术平台。第二届北大-清华统计论坛于 2017 年 6 月 26 日在清华大学成功举办。本次论坛由清华大学统计学研究中心及北京大学统计科学中心共同举办,共有来自海内外多所高校的 80 余名老师及同学参加了本次活动。2017 年 9 月,中心第三批四名博士研究生入学,中心在读研究生规模已达十名,这是对北大在统计学人才培养体量不足的一个有限的改进。

2017 年中心教员共发表 42 篇学术论文,其中有 4 篇发表在国际统计学四大顶级期刊。继 2016 年 3 月发布《空气质量评估报告(二)》,中心环境统计科研团队在三月份发布了《空气质量评估报告(三)---北京地区 2013-2016 年区域污染状况评估》,八月份该研究团队又发布了《空气质量评估报告(四)---京津冀 2013-2016 年区域污染状况评估》。这一系列报告引起了国内外媒体的关注,在社会上反响很大。中心在这一年中进一步同国内外的学者进行了有效的学术交流,共举办学术报告 45 场,其中来自海外的学者 30 场,国内的学者 15 场。这些学术报告拓宽了老师和学生的学术视野,促进了统计学领域的交流。我们衷心感谢所有对北京大学统计科学中心发展给予关心和指导的各级领导、各界朋友和海内外同行。



二、中心要闻

1. 北大统计学科在教育部第四轮学科评估中获评 A+



本一级学科中，全国具有“博士授权”的高校共 54 所，本次参评 51 所；部分具有“硕士授权”的高校也参加了评估；参评高校共计 120 所（注：评估结果相同的高校排序不分先后，按学校代码排列）。

评估结果	学校代码及名称
A+	10001 北京大学
	10002 中国人民大学
A	10055 南开大学
	10200 东北师范大学
	10269 华东师范大学
	10384 厦门大学

12 月 28 日，教育部学位与研究生教育发展中心公布了全国第四轮学科评估结果，在这一轮评估中，北京大学统计学一级学科获评 A+，位列全国高校榜首，另一所获评 A+ 的高校是中国人民大学。

学科评估是教育部学位与研究生教育发展中心按照国务院学位委员会和教育部颁布的《学位授予与人才培养学科目录》对全国具有博士或硕士学位授予权的一级学科开展整体水平评估。2017 年完成的第四轮评估在全国 95 个一级学科范围内开展(不含军事学门类等 16 个学科)，共有 513 个单位的 7449 个学科参评。评估结果按照“精准计算、分档呈现”的原则，根据“学科整体水平得分”的位次百分位，将前 70% 的学科分为 9 档公布，而 A+ 档表示该学科水平位列全国前 2%（或前 2 名）。在统计学一级学科中，全国具有“博士授权”的高校共 54 所，本次参评 51 所；部分具有“硕士授权”的高校也参加了评估；参评高校共计 120 所。2011 年 2 月，统计学成为理学门类下的一级学科；2012 年，

在教育部组织的第三次学科评估中，北京大学统计学科位列全国第二。

在本次评估中，北京大学共有 21 个一级学科获评 A+，统计学科是其中一员。北京大学这次学科评估中 A+ 学科数量和 A 类及以上学科占比均居全国高校之首。

北京大学的统计学科拥有雄厚的科研实力和光荣的历史底蕴，自 1940 年许宝騄先生从英国获统计学博士学位回国任教，首次在我国大学开设数理统计课程开始，几十年筚路蓝缕，北京大学为中国统计学科人才培养、学科发展做出了巨大的贡献。今天的北大统计站在了一个前所未有的高度，也必将不忘初心，砥砺前行，开创更辉煌的事业。

2. 北大统计学科入选“双一流”建设学科

2017 年 9 月 20 日，教育部、财政部、国家发展改革委印发《关于公布世界一流大学和一流学科建设高校及建设学科名单的通知》，公布了世界一流大学和一流学科（简称“双一流”）建设高校与建设学科名单。北京大学顺利进入一流大学建设高校 A 类名单（全国 36 所），北京大学统计学与其他 40 个学科入选一流学科建设名单。

在一流学科建设名单中，共有八所高校的统计学学科入选，它们是：北京大学、中国人民大学、清华大学（统计学与运筹学）、南开大学、东北师范大学、华东师范大学、上海财经大学、厦门大学。

北京大学入选的 41 个学科分别为：哲学、理论经济学、应用经济学、法学、政治学、社会学、马克思主义理论、心理学、中国语言文学、外国语言文学、考古学、中国史、世界史、数学、物理学、化学、地理学、地球物理学、地质学、生物学、生态学、统计学、力学、材料科学与工程、电子科学与技术、控制科学与工程、计算机科学与技术、环境科学与工程、

软件工程、基础医学、临床医学、口腔医学、公共卫生与预防医学、药学、护理学、艺术学理论、现代语言学、语言学、机械及航空航天和制造工程、商业与管理、社会政策与管理。

“双一流”建设学科名单

(按学校代码排序)

序号	学校名称	学科名称
1	北京大学	哲学、理论经济学、应用经济学、法学、政治学、社会学、马克思主义理论、心理学、中国语言文学、外国语言文学、考古学、中国史、世界史、数学、物理学、化学、地理学、地球物理学、地质学、生物学、生态学、统计学、力学、材料科学与工程、电子科学与技术、控制科学与工程、计算机科学与技术、环境科学与工程、软件工程、基础医学、临床医学、口腔医学、公共卫生与预防医学、药学、护理学、艺术学理论、现代语言学、语言学、机械及航空航天和制造工程、商业与管理、社会政策与管理

建设世界一流大学和一流学科，是党中央、国务院作出的重大战略决策。双一流名单的出台是教育部认真总结经验，加强系统谋划，加大改革力度，完善推进机制，坚持久久为功，统筹推进世界一流大学和一流学科建设，实现我国从高等教育大国到高等教育强国的历史性跨越所制定的发展方案。

根据《总体方案》和《实施办法》，“双一流”建设将以5年为一个周期，建设高校实行总量控制、开放竞争、动态调整。下一步，高校“双一流”建设进入实操阶段，工作重心将从凝练学科方向、编制建设方案转化到全面落实上来。

建设世界一流大学和一流学科，是党中央、国务院在新的历史时期，为提升我国教育发展水平、增强国家核心

竞争力、奠定长远发展基础，作出的重大战略决策。

3. 第二届大数据时代下的高维统计建模与分析研讨会顺利举行



5月14日至16日，第二届大数据时代下的高维统计建模与分析研讨会在光华管理学院成功举行，本届会议由统计科学中心、数量经济与数理金融教育部重点实验室和光华管理学院商务统计与经济计量系共同主办，光华管理学院承办。全球高校的学者们齐聚光华管理学院阿里巴巴报告厅，进行关于高维统计建模与分析的分享

和交流。会议为期3天，包括青年学者论坛、对话主编圆桌论坛以及高维统计研讨会。会议分为10个专题，来自哈佛大学、宾夕法尼亚大学、耶鲁大学、罗格斯大学、密歇根大学、哥伦比亚大学、北京大学、清华大学、中国人民大学、中国科学院、复旦大学、云南大学等全球高校的三十四名统计学者作了精彩的学术报告。来自国内近300名高校师生以及业内人士听取了本次研讨会。

北京大学高松副校长15日出席了高维统计研讨会，并为会议致辞。高松副校长代表北京大学向来自世界各地的高校学者及前来听取会议的师生表示热烈欢迎，并指出，北京大学致力于统计学科的未来发展，并将其作为数据科学和人工智能发展这类广泛计划的一部分，这是北京大学在下一阶段重要战略领域之一。

【青年学者论坛】

14 日, 青年学者论坛举行。会议联合主席、COPSS 奖获得者、宾夕法尼亚大学蔡天文教授与会议联合主席、北京大学统计科学中心联席主任陈松蹊教授为会议致辞, 他们向与会嘉宾和听众表示了热烈欢迎, 对本次会议促进学科发展、学术交流表达了期待。

北京大学统计科学中心林伟研究员、黄辉研究员、贾金柱研究员分别主持了专题报告。来自北京大学、清华大学、中国人民大学、南开大学、复旦大学、哥伦比亚大学、密歇根州立大学、内布拉斯加大学林肯分校等 14 名海内外青年学者报告了他们的最新研究。研究内容包括带错误发现率控制的大规模数据流动态变点检测、高维协方差矩阵估计的最优调节参数选择、带协变量的社区发现、基于惩罚加权计分函数的稀疏泊松回归等。

【对话主编 · 圆桌论坛】

14 日上午, 对话主编圆桌论坛举行。蔡天文教授主持了圆桌论坛, 来自国际统计界顶尖杂志《The Annals of Statistics》的两位联合主编——宾夕法尼亚大学 Edward I. George 教授和密歇根大学 Tailen Hsing 教授, 《Journal of the American Statistical Association: Theory and Methods》联合主编、罗格斯大学 Regina Liu 教授, 《Statistical Science》主编、罗格斯大学 Cun-Hui Zhang 教授参加了圆桌论坛, 他们围绕期刊投稿和论文写作分别表达了自己的看法。

Edward I. George 教授认为, 学生应树立远大目标并且追求卓越, 如果不能严格要求自己,



就永远不会达到理想的目标。同时, 他表示, 好的文章都是经历过几次拒绝才找到他该有的位置, 也不要轻易的放弃一个成功研究的线路。

Tailen Hsing 教授表示, 如果文章被拒绝, 也会详细地将拒绝的原因报告给每一位投稿人。他承诺, 一定会详细的看整篇文章的内容, 权衡审稿人意见, 听取每一位主编的建议才会做出最后的决定。

Cun-Hui Zhang 教授指出, Statistical Science 是一个独特的期刊, 它更加倾向现代统计, 特别喜欢发表综述类的文章。论文的难度技巧适中, 文章中简要地讲述将来研究的方向是锦上添花的。他鼓励年轻的学者多向 Statistical Science 投稿。

Regina Liu 教授认为, 写文章要知道读者期望的是什么, 难点在哪儿, 从而重点强调文章的结果。

在几位主编发言之后, 蔡天文教授做了总结, 一是尽可能将最好的文章展现出来, 不要将审稿过程看作润色自己文章的过程, 也不要期望审稿过程会给论文带来很好的改进意见。二是文章都会被拒绝很多次, 但是不要埋怨或责怪主编, 因为很多时候仅仅就是文章本身的问题。

在接下来提问环节中,主编们又耐心解答了大家对于审稿周期长、争议性的审稿意见以及统计学科发展趋势的问题。



【第二届高维统计研讨会】

15日至16日,宾夕法尼亚大学 Edward I. George 教授、澳大利亚科学院院士 Matt Wand 教授、耶鲁大学 Harrison Zhou 教授与 Heping Zhang 教授、罗格斯大学 Regina Liu 教授与 Cun-Hui Zhang 教授、密歇根大学

Ji Zhu 教授、威斯康星大学麦迪逊分校 Yazhen Wang 教授与 Ming Yuan 教授、哈佛大学 Tianxi Cai 教授、爱荷华州立大学 Dan Nettleton 教授、中国科学院王启华教授、云南大学唐年胜教授、北京大学王汉生教授等高维统计领域著名学者分别发表了学术报告。报告内容涉及基于针板套索的贝叶斯惩罚混合、高维数据可加回归的惩罚估计、基于边抽样的网络交叉验证、从序列图片中分析复杂基因表现型的统计挑战、高维解释变量的正交分解预处理、量子计算与统计等。

16日下午,本届研讨会顺利闭幕,会议联合主席蔡天文教授与陈松蹊教授祝贺会议成功举办,感谢所有报告人的到来,感谢大会组织者和参与会议筹备的工作人员及志愿者的服务,并对下一次会议的举办表示期待。

4. 北京大学统计科学中心 2017 年“全国统计学优秀大学生”夏令营成功举办

2017 年 5 月 26-28 日,由北京大学统计科学中心主办的 2017 年“全国统计学优秀大学生”夏令营成功举行。夏令营共邀请 25 名来自全国各大高校统计学及相关专业的优秀大学生参加。

5 月 26 日中午,夏令营营员前往燕东园 32 号小楼报到并领取材料。下午 3 时,夏令营笔试在理科一号楼 1560 教室开始,笔试内容涵盖数学、统计学、计量经济学等多学科基础知识,不同知识背景的同学各展所长,交出了满意的答卷。

5 月 27 日上午进行的是教员科研宣讲环节,统计科学中心联席主任陈松蹊教授首先向入营同学的到来表示欢迎,陈老师回顾了北京大学统计学科 70 年的发展历程,向大家展示了北大统计学专业深厚的历史底蕴、强大的师资力量和丰硕的科研成果。陈老师强调,北京大学统计科学中心始终致力于协调北大全校统计研究力量,促进统计学与其他学科的交叉与融合,建设世界一流的统计研究机构;更重要的是中心“建设一流的统计学科需要一流的学生”,希望同学们能够加入到北大统计科学中心团队中。



随后,北京大学统计科学中心的全职教员先后向大家介绍了各自的研究方向与最近的研究进展,台下同学们就自己关心的培养方式、学科前沿问题与老师们进行了深入的交流。

5月27日中午,统计科学中心的全体教员、在读博士生与参营同学们在勺园餐厅燕园厅进行了师生聚餐,席间大家互相交流学习生活,老师们向大家传授科研经验。师生频频互动,交流融洽,同学们也对未来的科研生涯有了更加明晰的了解与规划。

5月27日下午到5月28日上午进行了参营同学的自我展示环节。25位同学先后进行15分钟的英文自我介绍,将自己本科期间参与过的学习、科研经历以及对未来的研究计划做了详尽展示,并回答了中心老师提出的问题,师生之间的印象也随之加深。

5月28日下午进行了营员面试,进入复试的同学先后就自己在夏令营期间阅读的英文文献进行答辩。

为期三天的夏令营汇集了中国统计学专业优秀的青年学生,大家展示风采,互相交流,开阔眼界,也坚定了自己未来对统计学专业研究的信心。

5. 第十三届计量经济学理论与应用国际研讨会成功举行

6月13日至14日,第十三届计量经济学理论与应用国际研讨会在北大光华管理学院成功举行,本次会议由北京大学光华管理学院商务统计与经济计量系、北京大学统计科学中心、北京大学数量经济与数理金融教育部重点实验室共同主办。耶鲁大学 Xiaohong Chen 教授受邀在本届研讨会中做计量经济学理论专题报告(ET Lecture),哈佛大学 Elie Tamer 教授、华盛顿大学 Yanqin Fan 教授、印第安纳大学 Joon Park 教授和 Juan Carlos Escanciano 教授、波士顿学院 Zhijie Xiao 教授、新加坡管理大学 Liangjun Su 教授也受邀在会上发表学术报告。来自美国、澳大利亚、英国、新加坡、西班牙等以及国内北京、上海、厦门、台湾、香港等多所高校的统计学者和学生参加研讨会并发表报告。



会议主席陈松蹊教授



涂云东教授

13 日上午, 研讨会举行开幕仪式。光华管理学院商务统计与经济计量系涂云东教授主持, 会议主席、北京大学光华管理学院商务统计与经济计量系联合系主任陈松蹊教授为本次会议致辞。陈松蹊教授介绍了光华管理学院和商务统计与经济计量系的发展历史, 并希望与会学者、嘉宾共同支持并帮助光华在计量经济学研究和教学中取得更高的水平。他祝愿各位报告人能够在此平台中畅所欲言, 交流思想, 在光华度过愉快的时光。

耶鲁大学 Xiaohong Chen 教授带来了关于半非参数模型的学术文章。她的报告主要根据四篇论文 (其中两篇已经发表在世界顶级计量经济学期刊 *Econometrica* 和 *Journal of Econometrics*), 介绍了筛选拟似然比方法在半非参数模型中的前沿作用, 展现了该方法广泛的实用性。

13 日下午, 华盛顿大学 Yanqin Fan 教授发表了关于带辅助数据的矩等式模型中的部分模型识别和推断的报告, 报告主要在允许非可加可分的矩等式模型中提出了一种新的矩函数, 然后分别对三种模型进行了部分模型识别分析。

哈佛大学 Elie Tamer 教授带来关于多元离散选择模型的报告。报告主要介绍了在允许条件更弱情况下的动态多元离散选择模型, 分析了模型的识别定义集合, 然后对模型进行估计, 最后展现了该模型包含了多种已有的模型, 具有非常重要的应用价值。

14 日, 印第安纳大学 Joon Park 教授发表了报告《函数型动态数据中的计量分析》, 报告主要针对平稳和非平稳的函数型数据中的自回归模型进行估计, 都得到很好的性质, 并将模型应用于股票数据。

波士顿学院 Zhijie Xiao 教授发表了关于对带有条件异方差的时间序列数据的条件分位数的估计报告, 报告主要针对在时间序列数据中存在条件异方差时, 对条件分位数的进行了模型估计和统计推断, 展现了条件分位数在金融风险度量和收益率分布与预测等方面的应用。

印第安纳大学 Juan Carlos Escanciano 教授发表了题为《半参数模型的识别与费歇尔信息》的报告, 报告主要基于非参数和半参数模型中的统计信息, 针对半参数模型识别问题提出了一个新的系统性方法, 并得到了该方法良好的性质, 还说明了该方法能广泛地应用于现有许多非参数和半参数模型的识别问题中。

6. 第二届北大-清华统计论坛成功举办

2017 年 6 月 26 日, 第二届北大-清华统计论坛在清华大学学生职业发展指导中心报告厅成功举办。本次论坛由清华大学统计学研究中心、北京大学统计科学中心共同举办, 共吸引了来自海内外多所高校的 80 余名师生参加。

论坛开始, 清华大学统计学研究中心刘军教授发表了热情洋溢的开幕辞, 他指出, 北大-清华统计论坛的创办初衷就是充分发挥两所高校在中国统计学界的带头作用, 促进两校之间统计师生的交流合作, 加强两校学者的强强联合, 促进统计学科的发展。他希望, 通过这次论坛, 两校师生, 尤其是博士生同学之间能够建立深厚的友谊, 不断交流、不断沟通、互相鼓励、互相促进。

大会分别邀请了清华大学统计学研究中心兼职教授、哈佛大学生物统计系终身教授、系主任林希虹教授与北京大学统计科学中心联系主任、北京大学光华管理学院商务统计与经济计量系联合系主任陈松蹊教授做特邀报告。

首先, 林希虹教授做了题为 “Analysis of Massive Genome, Exposome and Phenome Data: Research and Training Opportunities” 的精彩报告。林教授在报告中谈到, 随着技术的不断进步, massive data, 例如全基因序列数据与各种基因序列数据变得越来越容易获得, 这些数据的获得也给数据分析方面的研究带来了机会和挑战。林希虹教授在报告中针对 massive data 可能带来的问题提出了新的处理方法: 包括使用 rare variant analysis 处理 whole-genome sequencing association, 用 integrative analysis 处理不同种类的基因数据, 用 causal mediation analysis 处理环境数据等。



茶歇过后, 陈松蹊教授围绕环境统计的最新进展, 做了题为 “Cautionary Tales on Air Quality Improvement in Beijing” 的报告。陈教授讲到, 根据官方统计数据显示, 2016 年北京市 PM2.5 排放量同比下降了 9.9%, 但是陈松蹊教授研究团队经过对北京市 36 个环境监测点过去四年的空气污染物数据与过去七年的气象数据进行分析, 结果显示, 2016 年北京市 PM2.5 的减排可能是由包括气象原因在内的多种因素叠加造成的, 也就是说, 北京市政策调控带来的污染控制成果可能没有官方宣称的 9.9% 之高, 我们需要进一步审慎评估政策带来的效应。

本次论坛还特别为不久前辞世的著名统计学家陈家鼎先生设立了纪念环节, 北京大学房祥忠教授、艾明要教授共同主持了纪念仪式。陈家鼎先生生前曾担任北京大学数学科学学院副院长、中国概率统计学会理事长、中国统计学会副会长等重要职务, 陈家鼎教授毕生致力于中国的科教事业, 在同行中享有极高声誉。房祥忠教授与艾明要教授先后回忆陈家鼎先生生前杰出的学术成就与感人的师德师风, 论坛现场师生感怀大师风范, 纷纷表达悼念之情。

在论坛最后的海报展示环节, 北大、清华两校参会的博士生将自己近期的研究成果以海报的形式展示在会场周边, 全场参会师生自由走动, 互相交流各自领域的最新进展, 论坛现场互动频繁, 气氛热烈。最终, 经过两校老师现场投票, 北京大学博士生苗旺荣获 “优秀毕业生” 奖, 清华大学李祺、北京大学王菲菲、梁德才分获 “优秀海报奖”。

本次论坛通过一个下午的交流活动使北大、清华两校师生之间加深了解、增进友谊, 两校统计学领域的研究也有了更加密切的联系, 会议取得了圆满成功。

7. 房祥忠教授当选中国现场统计研究会理事长

2017 年 11 月 4 日至 5 日, 中国现场统计研究会在浙江省杭州市召开第十届代表大会暨 2017 年学术年会。大会上, 与会代表对上一届学会工作进行了总结, 选举产生了新一届理

事会，确定了研究学会今后的发展方向和主要工作任务。本届代表大会选举北京大学房祥忠教授为新一届理事长。

中国现场统计研究会是为国家一级学会，现支撑单位是北京工业大学。学会下设试验设计分会质量分会，医药与生物统计分会，气象、水文、地质分会，统计调查分会，生存分析分会，工程概率统计分会，可靠性工程分会，国际标准与技术法规分会，教育统计与管理专业委员会，医药食品优化专业委员会，多元分析应用专业委员会，建设统计技术分会，资源与环境统计分会，安全性统计技术分会，统计综合评价研究分会，高维数据统计分会，空间统计学分会，计算统计分会等 19 个专业委员会，是国内著名的学术团体之一。



房祥忠，北京大学数学学院概率统计系教授，曾任中国数学会概率统计学会秘书长、全国统计方法应用标准化数据处理和解释技术分委员会主任委员、中国现场统计研究会常务理事、教育部统计学专业教学指导分委员会秘书长，全国电工电子可靠性与维修性标准化技术委员会（SAC/TC24）委员。房祥忠教授在《中国科学》、Biometrics 等学术期刊发表论文 30 余篇。主持多项国家自然科学基金项目和多项航天和国防科研项目。负责制修订过三项国家标准。获得 2011 年国防科技进步二等奖，2002 年北京市科技进步二等奖，1997 年航天部科技进步二等奖。

8. 陈松蹊教授获得 2017 年度高等学校科学研究优秀成果奖自然科学类一等奖

近日，教育部科技发展中心发布了《2017 年度高等学校科学研究优秀成果奖（科学技术）授奖项目公告》。北京大学统计中心联席主任陈松蹊教授获得自然科学奖一等奖，获奖项目为《高维数据统计推断方法》。

2017 年度高等学校科学研究优秀成果奖（科学技术）在同行专家通信评审的基础上，召开了专家评审委员会会议，评出候选项目和候选人，经高等学校科学研究优秀成果奖(科学技术)奖励委员会审定，教育部批准后予以公布。高等学校科学研究优秀成果奖（科学技术）分设自然科学奖、技术发明奖、科技进步奖和青年科学奖。其中，自然科学奖授予在基础研

陈松蹊教授《高维数据统计推断方法》项目以高维数据所派生的对多元统计方法进行更新、验证的需求为动力，以高维数据在金融风险管理、生物基因的应用为背景，建立了一个高维数据统计检验的理论体系，解决了一系列重要的数理统计问题，取得了具有国际影响力的创新成果。八篇代表性论文均发表在国际统计学顶级期刊上，相关工作在国内外被广泛引用，引发了系列的后续研究工作。

9. 姚方教授当选数理统计学会会士

由于在函数型数据分析领域所做出的奠基性和开创性的贡献，姚方教授 2017 年当选国际数理统计学会会士 Institute of Mathematical Statistics (IMS) Fellow。

姚方教授的主要研究包括无限维和高维空间的函数型数据分析，在实际应用中包括医学研究中的脑扫描图像的变化、经济金融研究中的投资交易数据以及经济指数变化、气象研究中污染温等水间变化对健康的影响等。从数学的角度，此类变量更适合被作为函数空间里的



数据来进行统计分析,这就需要建立新的统计模型,方法和理论,同时这也是近年来国际上统计学研究的一个热点方向。主要科研成果包括对稀疏类型的函数型数据提出了更有效的统计估计和推断的基于非参数和有效降维的方法和理论,从而建立了可以统一处理实际应用中观测到的稠密和稀疏型函数数据的统一框架,也开创了一个新的统计学研究方向。其研究成果在国际上有显著影响且被广泛引用,现任 JASA, Statistica Sinica, JCGS, JMA, CJS 期刊副主编,曾任 Annals of Statistics, Bernoulli, EJS, JSPI 期刊副主编。

10. 统计科学中心科研工作进展

统计科学中心成员近年来在统计研究和统计应用方面取得了一系列重要的、国际领先的成果和奖项。

1) 超高维数据方面

陈松蹊教授研究小组建立了一套高维统计检验的理论和方法,解决了一系列重要的数理统计问题,取得了具有国际影响力的创新成果,并在实际中得到了重要应用。

[Li and Chen (2012)]首次在一般超高维非参数数据下,提出了两样本高维协方差矩阵 Σ_1 和 Σ_2 是否相等的检验,即检验 $\Sigma_1 = \Sigma_2$ 。该检验普遍适用于超高维数据。这一检验提出后,在生物信息等领域得到了广泛应用。它同陈-秦均值检验一起是目前对两个高维样本分布差异性的主要检验方法。该论文是 WoS 高引论文。

[Zhong, Chen and Xu (2013)]考虑在更一般的高维数据分布,并允许不同数据维度具有弱相关性(所谓的 α -mixing 7 (α 混合)数据)。证明在这更一般的条件下,经过阈值处理的陈-秦检验的检验边界也达到最优检验边界。这一研究极大地提高了人们在挑战性最强的稀疏和弱信号检验问题上的认知,并给出了具有更高的检验效用的高维均值检验。

2) 函数型数据分析方面

姚方教授主要研究包括无限维和高维空间的函数型数据分析,例如函数主因子分析和各类函数回归,分类和聚类模型;复杂结构数据的模型选择和降维方法;对非独立相关型的纵向数据和生存分析的综合模型的研究。函数型数据分析的兴起主要是因为传统数理统计以向量型变量为研究和观测的对象,而实际应用中很多数据记录都是随时间和空间发生连续性变化的,例如环境研究中对空气污染的监测记录,医学研究中的脑扫描图像的变化等。现阶段的主要研究方向是把连续的函数型数据和具有高维,网络或者流形结构的模型及方法进行有机的结合,以及最新的机器学习和深度学习技术,提出创新的统计模型和解决相关的理论问题,从而可以有效的应用到大规模的数据处理中去。

由于在函数型数据分析领域所做出的奠基性和开创性的贡献,2014 年获得由加拿大统计学会和数学研究中心联合颁发的授予博士毕业 15 年内在加拿大做出突出贡献的统计学家的 CRM-SSC 奖,2017 年入选国际数理统计学会(Institute of Mathematical Statistics)的 Fellow。现任 Journal of the American Statistical Association, Statistica Sinica, Journal of Computational and Graphical Statistics, Journal of Multivariate Statistics, Canadian Journal of

Statistics 副主编；曾任 Annals of Statistics, Bernoulli, Electrical Journal of Statistics, Journal of Statistical Planning and Inference 副主编。

3) 因果推断方面

周晓华教授 2016 年当选为美国科学促进会(American Association for the Advancement of Science) 会士 (Fellow), 美国科学促进会是世界上最大的科学和工程学协会的联合体。美国科学促进会会士均在各自领域做出特别贡献, 由促进会指导委员会、或 3 个现任会士、或行政主任推荐, 并得到指导委员会大部份委员认可。2015《中国科学: 数学》杂志优秀论文奖。

周晓华教授开展了关于因果推断中的中介变量分析的有关研究。中介值分析 (mediation analysis) 是社会科学和医药科学进行因果推断的重要方法, 它主要帮助研究人员理解中介变量在因果关系中发挥的作用。传统的中介值分析方法是 Baron 和 Kenny 提出, 但这一方法需要进行一个很强的“sequential ignorability”假定才能保证因果关系。周晓华教授提出一个新的因果推断模型, 这一模型不但放松了原本“sequential ignorability”的强假设, 更能应用于多水平中间关系、多成分的中介变量等更复杂的情形。作为一种基于估计方程的方法, 该模型可以利用广义估计方程处理相关数据, 利用逆概率加权的方法处理缺失数据。

4) 网络数据研究方面

王汉生教授过去这五年 (2013-2017) 最主要的工作是对网络数据 (Network Data) 的理论研究, 以及相关的产业实践。在理论研究方面, 我们研究了网络拓扑结构的产生机制, 以及概率建模的可能性。在网络数据方面, 主要是研究了复杂相依关系同拓扑结构的参数化表达。共发表相关英文学术论文 29 篇, 获得国家基金委资助 300 余万。

5) 生物信息方面

席瑞斌研究员在过去几年中发展了一系列统计、生物信息方法, 很多都发表到了统计、生物信息的顶级杂志。例如在 Xia et al. (2017) 中, 席瑞斌和他的合作者发展了一个整合多个算法探测结构变异的方法, 显著提高了结构变异探测的敏感度、特异性及断点准确性; 在 Xi et al. (2017) 中, 他们发展了一个新的基于全基因组测序数据探测拷贝数变异的算法, 并将此算法应用到了约 1000 个癌症病人的正常基因组上, 发现了一些包括 GOLPH3 在内的关键的癌症易感变异位点; Yuan et al. (2017) 则发展了一个用来探测网络差异的方法, 此方法主要基于一个新的 D-trace 凸损失函数, 证明了在超高维数下此方法的相合性。另外, 席瑞斌还和他的合作者一起做了很多基因组、癌基因组的大数据分析研究, 这些文章很多也都发表到了顶级杂志上。在 Chang et al. (2017) 中, 他们分析了 ~100 对食管癌基因组, 发现与饮酒相关的点突变指纹, 并找到了一些包括 EGFR 等基因在内的关键可靶向用药变异; 在 Gao et al. (2017) 中, 席瑞斌和他的合作者分析了肝癌基因组的异质性, 发现肝癌细胞有极高的异质性, 构建了相应的癌细胞进化树, 并发现癌细胞的异质性显著的影响了癌细胞对药物的响应; Kim et al. (2013) 则对 8000 对个癌基因组进行了拷贝数变异大数据分析, 发现了一些泛癌基因组中的关键拷贝数变异位点, 并研究了染色体碎裂现象的多种癌症中的基本特征; 最后, De Los Angeles et al. (2015) 系统研究了干细胞研究中的关键检验准则, 提出并展示了基因组分析在干细胞研究中的关键作用。

6) 计量经济和股指预测方面

在金融领域, 股指预测一直是行业的重点。涂云东博士发现股指预测中的显著非线性关系, 并结合金融学中的常见约束关系, 构建了带约束的非线性股指预测模型。该模型的预测表现比文献中的现有方法有较大改善。该结果形成两篇文章 Li, Tu and Ullah (2014, 2015), 分别发表在国际一流杂志 *Journal of Econometrics* 和 *Journal of Business and Economic Statistics* 上。

涂云东博士针对诺奖获得者 Engle and Granger 1987 年在 *Econometrica* 上提出的误差修正模型, 提出了模型选择和平均的策略, 改善了其在预测中的表现。该成果形成一篇文章, Tu and Yi (2017), 发表在国际一流杂志 *Journal of Econometrics*。

宋晓军博士与 bderrahim Taamouti 教授提出利用非参数方法来度量和检验时间序列中可能存在的非线性和线性格兰杰因果关系。另外, 该方法由于不假设任何参数模型, 所以并不会受到模型错误设定的影响。该文章于 2016 年发表在 *Journal of Business & Economic Statistics*。

7) 高维建模理论及其应用

贾金柱副研究员针对高位稀疏图模型, 创新性地设计了 MCMC 算法, 该方法可以有效地对高维稀疏图模型进行抽样 (Yangbo He, Jinzhu Jia and Bin Yu 2013, *Annals of Statistics*)。使用该 MCMC 算法, 我们研究了如何计算因果图中等价类的数量 (Yangbo He, Jinzhu Jia and Bin Yu 2015, *Journal of Machine Learning Research*)。我们研究了非齐次模型对于变量选择方法的影响, 理论上保证了变量选择方法的广泛适用性 (Jinzhu Jia, Karl Rohe and Bin Yu 2013, *Statistica Sinica*)。我们创新性地提出了预处理方法, 该方法可以更好地解决变量选择问题 (Jinzhu Jia and Karl Rohe 2015, *Electronic Journal of Statistics*)。我们应用稀疏学习方法很好地分析了文本新闻数据, 得到有意义的结果, 为社会学家分析新闻数据提供了很大方便 (Jinzhu Jia, Luke Miratrix, Bin Yu, Brian Gawalt, Laurent El Ghaoui, Luke Barnesmoore and Sophie Clavier 2014, *The Annals of Applied Statistics*)。



三、人员组成

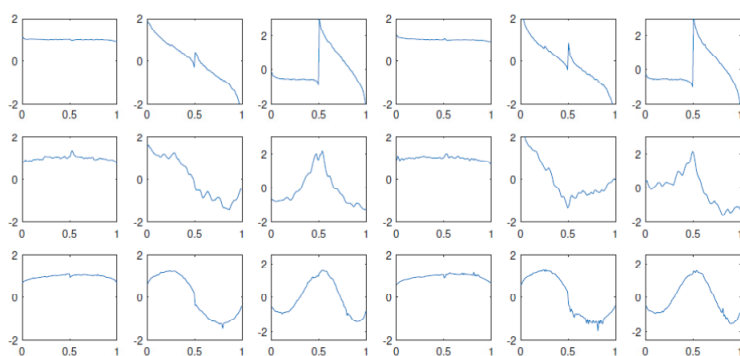
1. 人才引进



黄辉研究员：

统计中心与数学学院合聘。2010年在马里兰大学获统计学博士学位，2012-2013年在迈阿密大学管理科学系进行博士后研究工作，2014年加入北大统计中心和数学学院。2015年入选学“国家级重大人才青年计划”。研究方向：函数型数据分析，纵向数据分系，空间统计学，时间序列与空间点模式分系。

研究介绍：



在行为研究、金融研究或者社交网络研究当中，我们很多时候观测的数据都是事件数据，例如一个人是否参加会议，一支股票是买还是卖，朋友圈的段子是否转发，等等，都可以定义成我们感兴趣的事件。如果这类数据在时间轴上或者

空间上随机出现，那么针对其随机机制的研究就称为点模式分析。通常我们会假设点模式是由背后的点过程决定的，最常见也是最简单的点过程是泊松过程。如果进一步，泊松过程强度函数（intensity function）的随机性由一个或者多个高斯过程来解释，这类点过程就称为 Log-Gaussian Cox 过程。我们近期投稿至 JRSSB 的一篇文章中，尝试了一种为 Log-Gaussian Cox 过程构造多水平、半参数模型的建模框架。该模型受到了我国股票交易数据的启发。我们拥有上海股票交易所上百万个账户近两年半每天的股票交易行为数据。由于股票的买卖行为只能发生在工作日早上九点至下午三点，而不同的交易日之间又存在一定的相依关系，因此该数据至少具有两层水平：在日期（date）水平上，每个账户的交易行为可看成离散时间的时间序列，而所有个体又在独立地重复某种买卖行为特征。除此以外，个体和日期之间还有交互作用。我们利用多水平函数型主成分分析技术，对每个水平对应的隐藏高斯过程进行分解，同时为了获得高斯过程协方差核（covariance kernel）的相合估计，我们使用了条件似然函数（conditional likelihood）。该方法不仅可以应用到单变量的多水平点过程上，也可自然地拓展到多变量情形。例如在股票交易数据中，我们对买行为和卖行为进行了联合的统计建模（joint modeling），发现了买行为的特征函数（下图左三列）和卖行为的特征函数（右三列）非常相似。



贾金柱研究员：统计中心与医学院公卫学院合聘。2009 年北京大学概率统计系博士，2009-2010 年在 UC Berkeley 统计系进行博士后研究工作，2011 年加入统计中心。研究方向：高维统计推断、统计机器学习和因果推断等。

研究介绍：

判别分析是一个经典的统计学和机器学习问题。在模式识别、生物信息学中有广泛的应用。LDA (Linear Discriminant Analysis) 是应用广泛的判别分析方法。LDA 的原理是将带上标签的数据点投影到维度更低的空间中，使得投影后的点会形成按类别区分，相同类别的点将会在投影后的空间中更接近。经典统计教科书告诉我们，在一些条件下 Fisher 线性判别分析(LDA)是最优的。然而当维度很高时，LDA 的表现并不好。如 Bickel and Levina (2004) 指出在高维情形 LDA 甚至不如随机猜测。

问题出现在高维协方差矩阵的估计上，因为要估计的参数非常多，高维协方差矩阵出现了很大的误差，甚至出现协方差阵不可逆的情况。为了解决这一难题，统计学家提出修正的协方差估计，并证明了许多优良的性质。也有统计学家指出可以通过规避估计协方差来解决这一难题[CL2011]。其实，Fisher 最早建立了判别分析和回归分析之间的联系 [F1936]：二分类的 LDA 等价于线性最小二乘方法——即将两个类别分别用 0 和 1 表示，然后求解线性最小二乘。有了这一联系，带有变量选择的高维线性判别就转化成了传统的 Lasso 求解问题。虽然这一转化很自然，其理论结果有待进一步探讨。原因是 Lasso 的理论结果局限于线性模型，其误差的高斯型尾概率分布起决定作用。LDA 虽然可以转化为类似线性回归的最小二乘方法，但是其本质的模型已经很不一样。Mai et al(2012) 建立了 Lasso 求解 LDA 所得模型选择相合性的结果。通过细致分析，我们发现使用 random matrix 理论，可以证明，如果起决定作用的变量个数很少时，Lasso 可以得到最优的判别结果，这一点和低维经典 LDA 结论类似。

参考文献：

Yanfang Li and **Jinzu Jia**.(2017) L1 least squares for sparse high-dimensional LDA. *Electronic Journal of Statistics*, 2017, 11(1):2499-2518.



李程高级研究员：统计中心与生命科学学院合聘。2013 年加入统计中心。研究方向：生物信息学和计算生物学，癌症基因组学和个体化医疗，高通量生物数据分析中的统计建模。目标为使用计算生物和基因组学来帮助预测、发现、验证新的癌症基因和机理，从而为新药物和个体化医疗作出贡献。

研究介绍：

随着三维基因组技术的快速发展以及它的广泛应用前景，美国国立卫生研究院在 2014 年制定了 4D Nucleome 计划，资助多个研究团队从三维空间和时间尺度上研究细胞核内染色质的组织结构和功能，体现了该领域的重要性和前沿性。我们基于对癌症基因组中非整倍体变异频繁出现的原因和后果的研究兴趣和基础，在 2014 年开始将研究重心聚焦在癌症三维基因组学领域，通过自主建立的 Hi-C 实验和分析流程，首先研究多发性骨髓瘤细胞中非整倍体变异对三维基因组和表达谱的影响。在过去四年间，我们在三维基因组学实验技术开发、分析算法、数据库网站、与临床研究团队开展合作方面都取得了重要进展。

近期发表工作：

1. Wu P, Li T, Li R, Jia L, Zhu P, Liu Y, Chen Q, Tang D, Yu Y, **Li C.** (2017) 3D genome of multiple myeloma reveals spatial genome disorganization associated with copy number variations. *Nat Commun.* 8(1):1937.
2. Hydbring P, Wang Y, Fassl A, Li X, Matia V, Otto T, Choi YJ, Sweeney KE, Suski JM, Yin H, Bogorad RL, Goel S, Yuzugullu H, Kauffman KJ, Yang J, Jin C, Li Y, Floris D, Swanson R, Ng K, Sicinska E, Anders L, Zhao JJ, Polyak K, Anderson DG, **Li C**, Sicinski P. (2017) Cell-Cycle-Targeting MicroRNAs as Therapeutic Tools against Refractory Cancers. *Cancer Cell.* 31(4):576-590.e8.
3. Wang Y, Fan C, Zheng Y, **Li C.** (2017) Dynamic chromatin accessibility modeled by Markov process of randomly-moving molecules in the 3D genome. *Nucleic Acids Res.* 45(10):e85.
4. Li R, Liu Y, Li T, **Li C.** (2016) 3Disease Browser: A Web server for integrating 3D genome and disease-associated chromosome rearrangement data. *Sci Rep.* 6:34651.



林伟研究员：统计中心与数学学院合聘。2011年在南加州大学获得应用数学（统计方向）博士学位, 2011-2014年在宾夕法尼亚大学生物统计与流行病学系进行博士后研究工作, 2014年加入北大统计中心和数学学院, 2015年入选“国家级重大人才青年计划”。他的研究方向主要是高维统计和大数据建模与分析, 此外还包括因果分析、生存分析、成分数据分析、时空统计等, 及其在基因组学、宏基因组学和环境科学中的应用。

研究介绍:

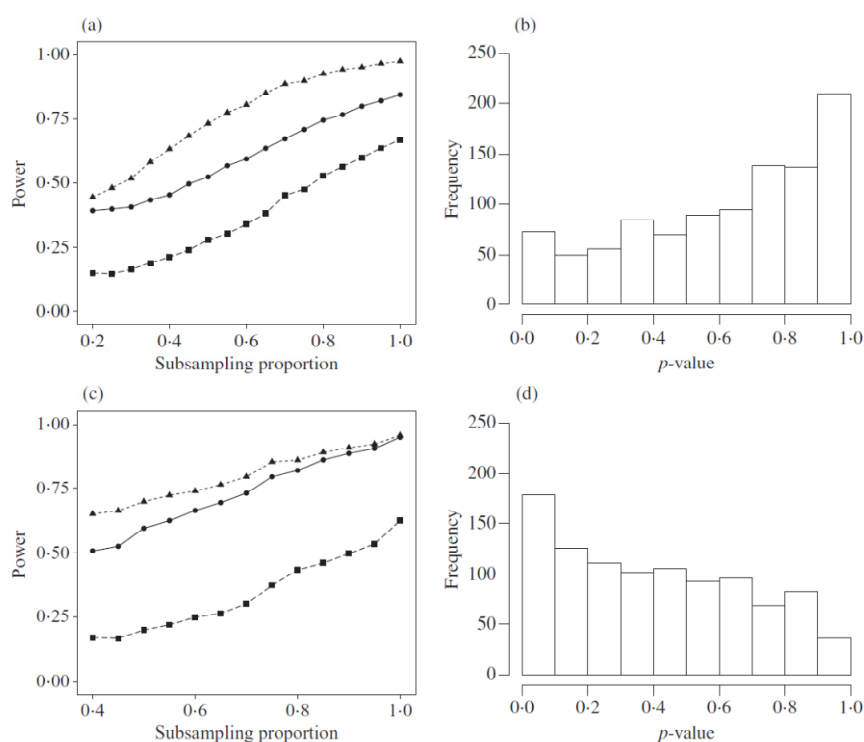


Fig. 1. Analysis of two microbiome datasets: empirical power curves of the proposed test (triangles) and the tests based on log-transformed (circles) and raw (squares) compositions with $\alpha = 0.05$ are shown in (a) for the obesity data and (c) for the Crohn's disease data; histograms of p -values from the proposed test in the back-testing are shown in (b) for the obesity data and (d) for the Crohn's disease data, for 1000 replicates.

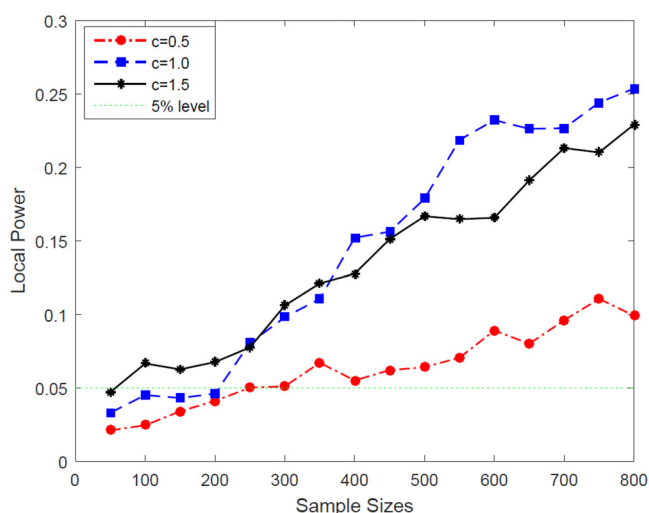
林伟研究员及合作者在 *Biometrika* 上发表论文, 针对高维成分数据提出一种两样本均值检验方法, 为基于成分数据的科学研究和数据分析增添了新的研究手段。成分数据是指各分量为正且加和为一的数据, 在宏基因组学中测量微生物组成、经济学中描述家庭支出构成、文本分析中刻画词频分布等应用中都十分常见。成分数据的特点导致其各分量之间复杂的相依关系, 造成许多经典统计方法的失效甚至得到误导性的结果。该论文在两样本均值检验的一般框架下, 首先定义了一种可检验的成分等价性假设, 然后基于成分数据的中心化对数比变换发展了一种成分等价性检验方法, 并推导了检验统计量的渐近分布及功效。论文将该方法用于肥胖症和克隆氏症中的肠道微生物组成成分检验问题, 取得了良好的效果。



宋晓军研究员：统计中心与光华管理学院合聘。2014 年西班牙马德里卡洛斯三世大学经济系经济学博士，同年加入北大统计中心。研究方向：计量经济学理论（非参数与半参数方法，模型设定检验，自助方法，时间序列分析和实证资产定价）及应用计量经济学。

研究介绍：

宋晓军研究员及合作者撰写论文(*Journal of Econometrics*,已接受)，针对条件分布函数的对称性检验提出了一个新方法。由于条件对称性假设在计量经济学和统计学上的重要性，如何有效的检验该假设至关重要。本文提出的方法不仅适用于截面数据也适用于时间序列数据，并且能够有效的避免模型误设的问题。传统检验在验证条件对称性时，通常需要假设对称中心是已知的常数或者函数形式，而当函数形式并不是所假设的形式时，基于该参数假设所构建的检验显然会出现检验水平和能效方面的问题。本文利用非参数方法估计对称中心，然后基于非参数残差构建统计量。另外由于该论文运用了条件特征函数来检验对称性，所以并不需要假设残差是独立于解释变量的。这样构建检验统计量的话，就允许高阶相依性的存在，



而这正是一般金融时间序列的特征。该论文进一步推导了检验统计量的大样本性质，提出一个乘数自助法来实现该检验方法。大量蒙特卡洛模拟表明该方法能够很好的保持水平，并且对各种各样的非对称性备择假设都具有较好的能效。最后，论文将该检验方法用于检验四个主要股票市

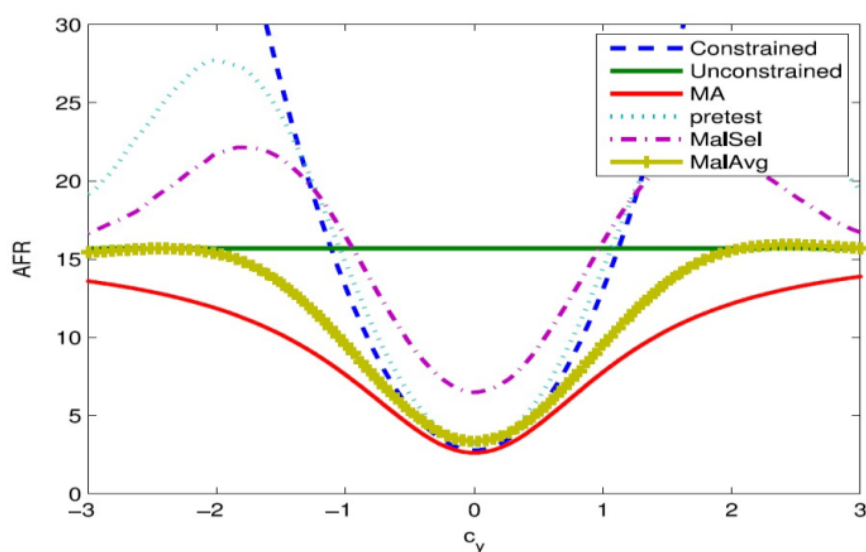
场日回报率在 2008 年经济危机前后的表现，以及检验商业周期是否对称等问题上，取得良好的实证结果。此外，本论文所发展的技术还可以检验其它非参数类型的约束，例如在数据具有相依性情况下的条件相关性，条件独立性检验等。



涂云东研究员：统计中心与光华管理学院合聘。2012 年加州大学经济学博士，同年加入统计中心。研究方向：非参数建模，高维数据建模，信息计量经济学，金融计量学，经济预测，生产效率建模，宏观计量经济学等。

研究介绍：

涂云东研究员和他的合作者 2017 年在 *Journal of Econometrics* 上发表论文 (Tu and Yi, 2017, JoE)，利用模型选择和平均的方法，来解决协整误差修正模型中如何利用误差修正项的关键问题。诺奖得主 Robert Engle 和 Clive Granger 于 1987 年在 *Econometrica* 上发表文章，提出了非平稳单位根过程之间存在协整，即单位根过程的线性组合成为平稳的序列，以及其对应的误差修正模型 (Engle and Granger, 1987)。在协整的诸多表述中，误差修正模型尤为重要，特别是对于理解经济结构和预测来讲。因为它描述了不平稳变量之间的长期均衡，即协整关系，是如何通过误差修正项来影响它们的短期波动的（差分变量）。因此，误差修正模型被用来构建对非平稳协整变量的预测。然而，研究表明，误差修正项的系数往往在统计上是不显著的，由此所导出的预测也并不是总比简单的差分模型好 (Hoffman and Rasche, 1996; Jansen and Wang, 2006)。如何利用误差修正项来构建预测成为研究的关键问题。Tu and Yi (2017) 首次在协整的误差修正模型中讨论使用模型选择和平均的方法来解决如何利用误差修正项的问题，并且建立了模型选择和模型平均预测的风险渐近表达式，提出了可行的模型选择和模型平均估计量，通过数值模拟和应用实例说明模型平均方法带来的预测风险更小。





席瑞斌研究员:统计中心与数学学院合聘。2009年华盛顿大学数学博士，2012 年加入北大统计中心和数学科学学院，2013 年入选“国家级重大人才青年计划”。席瑞斌博士的研究方向是生物数据，特别是高通量生物数据的统计计算和分析，大数据及高维数据的分析以及贝叶斯统计。

研究介绍:

结构变异是指会导致基因组结构发生变化的基因组变异。结构变异在正常人基因组中广泛存在，并可能与包括癌症在内的多种疾病相关。目前探测结构变异主要基于第二代测序数据，研究者已经发展了多个基于二代测序数据探测结构变异的方法。然而，这些方法的特异性及敏感性都有待进一步提高，不同算法所预测的结构变异常常有很大区别，因此造成后续分析极大的困难。席瑞斌发展了一个探测结构变异的算法 SVmine，此算法通过整合多个算法预测结果从而提高探测结构变异的能力。大量的模拟和实际数据分析表明，无论是结构变异检测的特异性、敏感性还是断点的准确性，SVmine 均显著好于其他算法（图 1）。此文发表于 Bioinformatics 杂志。Bioinformatics 是由牛津大学出版社出版的生物信息学杂志，主要发表生物信息及计算生物方面的学术论文，是国际公认的生物信息学权威杂志，2016 年影响因子 7.3。北京大学博雅博士后夏禹超是论文的第一作者，博士生刘芸为该文做出了重要贡献。

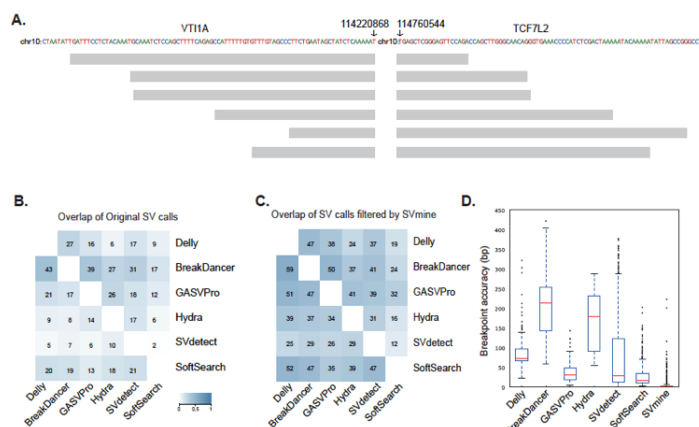


图 1. A. SVmine 在一个结直肠癌基因组中探测到的基因融合 VTI1A-TCF7L2, 其预测的断点位置与 Sanger 测序验证的断点位置完全相同。B. 在 8 个结直肠癌基因组中，6 个算法的体细胞结构变异预测的两两相交情况，图中的数字表示横轴和纵轴算法共同预测的结构变异在横轴算法预测结构变异中的比例。例如，一行二列的数字 27 代表 27% 的 Delly 预测的结构变异同时也在 BreakDancer 预测的结构变异中。C. 经过 SVmine 整合分析后，算法两两预测交集有显著的提高。图中的数字与 B 类似。例如，经过 SVmine 整合后，现在 47% Delly 预测的结构变异同时也在 BreakDancer 预测的结构变异中。D. 不同算法预测结构变异的断点位置与 Sanger 测序验证断点位置的箱式图，SVmine 所预测的断点位置在所有算法中最准确。

2. 中心部分成员

姓名	工作	研究方向	学校院系
陈松蹊	联席主任, 执行委员	金融计量学, 经济计量学, 风险度量, 随机过程, 统计推断, 高维数据分析, 抽样方法	光华管理学院商务统计与经济计量系
耿直	联席主任, 执行委员	因果推断, 数理统计, 生物医学统计, 因果网络, 贝叶斯网络	数学科学学院概率统计系
姚方	副主任, 执行委员	函数型数据分析, 复杂结构数据, 高维数据, 非参数回归	数学科学学院概率统计系
李程	执行委员, 中心全职教员	生物信息学, 癌症基因组学, 大数据统计建模	生命科学学院生物信息专业
王汉生	执行委员	高维数据分析, 变量选择, 数据降维, 值理论, 以及半参数模型	光华管理学院商务统计与经济计量系
吴岚	执行委员	金融资产定价, 金融统计, 经济资本模型, 精算学	数学科学学院金融数学系
任艳霞	执行委员(自七月)	测度值马氏过程、分支粒子系统等	数学科学学院概率统计系
周晓华	执行委员(自七月)	缺失数据, 因果推断分析, 大数据分析, 半参数模型, 医学检验, 脑科学, 卫生经济, 卫生服务领域发展新的统计方法	北京国际数学研究中心
黄辉	中心全职教员 (八月份离职)	函数型数据分析, 空间统计学	数学科学学院概率统计系
贾金柱	中心全职教员	高维统计推断, 统计机器学习, 因果推断	数学科学学院概率统计系
涂云东	中心全职教员	金融计量理论, 理论计量经济学, 应用计量经济学	光华管理学院商务统计与经济计量系
席瑞斌	中心全职教员 讲座召集人(下半年)	统计计算, 生物信息	数学科学学院概率统计系
林伟	中心全职教员 讲座召集人(上半年)	高维统计, 大数据问题, 因果推断, 生存分析	数学科学学院概率统计系
宋晓军	中心全职教员	计量经济学理论及应用计量经济学	光华管理学院商务统计与经济计量系
张志华	中心教员	机器学习与应用统计学	数学科学学院概率统计系
陈嵘	特聘教授	非线性和多变量时间序列分析, 蒙特卡洛方法, 统计计算和贝叶斯分析及其应用	Rutgers University
金加顺	特聘教授	大规模多元假设检验, 癌症分类, 变量选择, 谱聚类和主成分分析 (PCA), 图论和精度矩阵, 网络分析, 随机矩阵理论	Carnegie Mellon University
邓明华	统计咨询召集人	生物信息学, 计算生物学	数学科学学院概率统计系
张俊妮	统计咨询召集人 讲座召集人	因果推断, 贝叶斯分析, 蒙特卡洛方法, 数据挖掘及其应用	光华管理学院商务统计与经济计量系
房祥忠	政府统计工作室召集人 (执行委员, 至七月)	生存分析, 可靠性, 应用统计	数学科学学院概率统计系
虞吉海	政府统计工作室召集人	空间计量, 面板数据	光华管理学院商务统计与经济计量系

四、科研项目

序号	项目名称	负责人	总经费 (万元)	项目来源	类别	开始时间	结题时间
1	生物数据统计分析的方法、理论与应用	艾明要	48.0	基金委	协作项目	2014.01	2018.12
2	试验设计的统计推断基础理论	艾明要	48.0	基金委	面上项目	2017.01	2020.12
3	高维向量的分类技术研究合同	艾明要		企事业单位	技术服务	2016.12	2017.03
4	大数据驱动的管理决策模型与算法	陈松蹊	266.0	基金委	重点项目	2016.01	2020.12
5	空气质量统计评估方法	陈松蹊	414.6	科技部	重点研发计划	2016.07	2020.06
6	金融连续时间随机过程的统计推断	陈松蹊	57.5	基金委	面上项目	2014.01	2017.12
7	空气质量分析测算技术服务	陈松蹊		中国环境监测总站	合作合作	2017.1	2018.1
8	2013-2017 年环境空气质量监测数据统计	陈松蹊		中华人民共和国环境保护部	技术合作	2017.11	2018.11
9	市场冲击模型中引入执行风险后之最优执行策略及相关问题的研究	程雪	18.0	基金委	青年基金	2017.01	2019.12
10	与 Levy 过程驱动的倒向随机微分方程相关的随机控制和金融问题	程雪	9.0	基金委	协作项目	2015.01	2018.12
11	基于网络的全基因组关联分析	邓明华	70.0	基金委	面上项目	2015.01	2018.12
12	疾病相关生物分子网络的构建、动态演化及调控规律（中科院生物物理所）	邓明华	285.0	科技部	重大研究计划	2015.01	2019.08
13	生物数据统计分析的方法、理论与应用	耿直	48.0	基金委	协作项目	2014.01	2018.12
14	食品安全统计模型与数据分析研究	耿直		其它部门	其它专项	2016.06	2017.05
15	多模态脑语言、运动功能图谱构建及其在脑胶质瘤手术中的应用	耿直	30.0	科技部	国家 863 计划	2015.01	2017.12
16	技术合作协议	耿直		企事业单位	技术合作	2014.04	2017.03
17	高维图模型的结构空间及学习	何洋波	48.0	基金委	面上项目	2017.01	2020.12
18	国家级重大人才青年计划科研启动费	黄辉	16.0	教育部	启动经费	2015.01	2017.12

动费

19	大气污染物的时-空统计模型	黄辉	282.0	科技部	重点研发计划	2017.01	2021.12
20	基于函数型数据分析的联合统计建模：理论与应用	黄辉	22.0	基金委	青年基金	2015.01	2017.12
21	文本挖掘的统计建模	贾金柱	45.0	基金委	面上项目	2016.01	2019.12
22	金融衍生品定价中的计量经济学分析	李辰旭	48.0	基金委	重点研发计划	2017.01	2020.12
23	税务，人社大数据分析项目	李东风		企事业单位	技术服务	2016.05	2017.05
24	国家级重大人才青年计划科研启动费	林伟	16.0	教育部	启动经费	2015.01	2017.12
25	高维复杂数据的稀疏与低秩建模及推断	林伟	48.0	基金委	面上项目	2017.01	2020.12
26	基于统计与数值模式的多污染物数据场构建	林伟	317.3	科技部	重点研发计划	2017.01	2021.12
27	无穷维随机微分方程遍历理论及其相关课题中的一些问题	刘勇	55.0	基金委	面上项目	2014.01	2017.12
28	超过程的极限理论	任艳霞	48.0	基金委	面上项目	2017.01	2020.12
29	高维非线性模型中的因子分析：理论和应用	涂云东	46.4	基金委	面上项目	2017.01	2020.12
30	高维复杂数据的理论与应用	王汉生	245.0	基金委	杰出青年基金	2016.01	2020.12
31	大气污染数据质量监控	王汉生	233.5	科技部	重点研发计划	2016.07	2020.06
32	北京大学-华农天时共建“北京大学农业大数据研究中心”	吴岚		企事业单位	技术服务	2016.12	2026.12
33	保险业信息披露监管制度研究	吴岚		其它部门	其它专项	2015.09	2016.06
34	国家级重大人才青年计划科研启动费	席瑞斌	70.0	教育部	启动经费	2015.01	2017.12
35	基于高通量测序数据研究基因组变异的统计问题	席瑞斌	60.0	基金委	面上项目	2015.01	2018.12
36	金融中的动态 copula 理论及其应用研究	杨静平	48.0	基金委	面上项目	2017.01	2020.12
37	污染与宏观经济、人口健康的计量分析	虞吉海	355.6	科技部	重点研发计划	201607	202006
38	排他过程中的若干问题	章复熹	50.0	基金委	面上项目	2014.01	2017.12
39	金融不良资产估值定价模型研究	张俊妮		中国信达资产管理股份有限公司	技术合作	2014	2017

五、博士毕业生去向

2017 年，北京大学统计及金融数学专业共有 19 名博士毕业生，毕业时 18 名已落实工作，其中 9 人在国内外知名高校获得教职或从事博士后工作，9 人在知名互联网企业或银行等大型金融机构工作。

姓名	专业	工作单位	指导教师
白 洋	应用数学（金融）	普信恒业科技发展（北京）有限公司	吴 岚
韩世予	应用数学（金融）	华夏基金管理有限公司	吴 岚
黄晨笛	概率论与数理统计	毕业时未定	姚 远
孔祥顺	概率论与数理统计	北京理工大学	艾明要
李艳芳	概率论与数理统计	北京百分点信息科技有限公司	耿 直
李照男	概率论与数理统计	中国工商银行股份有限公司金融市场部	房祥忠
梁 萱	统计学	Monash University 博士后	陈松蹊
林 锋	应用数学（金融）	中国银行股份有限公司	杨静平
苗 旺	概率论与数理统计	哈佛大学博士后	耿 直
邵 辉	应用数学（金融）	新加坡国立大学博士后	杨静平
王 健	统计学	全球能源互联网发展合作组织	张俊妮
王 莹	统计学	奥克兰大学博士后	陈松蹊 涂云东
王菲菲	统计学	中国人民大学博士后	张俊妮
熊杰超	概率论与数理统计	腾讯科技（深圳）有限公司	姚 远
殷云剑	概率论与数理统计	银华基金管理股份有限公司	周晓华
俞翰君	概率论与数理统计	首都经济贸易大学	房祥忠
袁慧莉	概率论与数理统计	中国光大银行股份有限公司	邓明华
赵琬迪	统计学	首都经济贸易大学	王明进
朱雪宁	统计学	宾州州立大学博士后	王汉生



六、交流访问

1. 新西兰统计局资深研究员 John Bryant 访问统计科学中心

2017年2月6日至3月12日,新西兰统计局资深研究员 John Bryant 访问了统计科学中心。John Bryant 是人口学博士,在新西兰统计局推广用贝叶斯方法估计和预测小区域特征,在人口学界具有国际知名度。

小区域是指按照性别、年龄和地区等因素划分的子群体。估计和预测小区域特征而不仅仅是总体特征,能更切合政策制定者和行业决策者的需求。对于总体特征而言,样本通常很大因而可以进行直接估计;但是对于比较细致的小区域,样本量就比较小,所以无法直接采用属于这些小区域的样本对其特征进行精确的估计。需要通过统计建模从其他区域或时期的数据借力才能更好地估计小区域特征。

John Bryant 之前曾多次访问统计科学中心并与张俊妮教授合作。他们发展了一套小区域估计和预测的贝叶斯框架,能够以一致连贯的方式应对不确定性和缺失值,很好地结合不同来源的数据,并很好地引入来自专家判断的信息。此次访问中,John Bryant 与张俊妮教授合作,详细讨论了他们正在合作撰写的《Bayesian Demographic Estimation and Forecasting》一书(将由 CRC Press 出版)的整体框架,并进行了一些章节的写作。他们还与统计科学中心的博士生王菲菲合作,研究了如何使用贝叶斯非参数模型刻画小区域划分因素之间复杂并随时间变化的交互效应,并完成了相关程序。

John Bryant 的访问进一步加深了他与统计科学中心的合作,也体现了统计科学中心对国际合作研究的大力支持。

2. 普林斯顿大学 Yacine Aït-Sahalia 教授讲座成功举办



3月10日,应北京大学统计科学中心邀请,世界著名的计量金融专家,普林斯顿大学的 Yacine Aït-Sahalia 教授访问北大。Yacine Aït-Sahalia 教授是普林斯顿大学著名的跨学科研究中心 Bendheim 金融中心教授,曾担任《金融研究评论》和《计量经济学杂志》等著名的国际统计学与计量金融学期刊的联合编辑。

Yacine Aït-Sahalia 教授在光华一号楼 112 教室发表了题为“连续时间模型中转移密度的估计问题”的精彩演讲。该演讲由李辰旭老师主持,北大师生近百人参加。

讲座伊始, Yacine 教授回顾了自己与北大的渊源, 迅速拉近了与听众的距离。随后, 以大家熟知的随机微分方程引出了本次讲座的主题。Yacine 教授通过回顾极大似然估计原理, 指出离散频率样本情形下, 使用极大似然估计法进行估计往往是简便的。借鉴这一思路, 将问题推广到连续时间情形, 考虑一个 binomial tree 的模型进行逼近, 我们便可以得到转移密度显示表达式的一个高度精确的近似。随后, Yacine 教授以单变量为例, 将这一过程中的具体问题与有关性质予以详细的数理推导, 同时, Yacine 教授将自己的数据模拟结果也与大家进行了分享。这一方法的提出, 对连续时间变量转移密度形式的近似估计这一金融学与统计学颇为关注的传统领域提供了新的思考方向。

在问答环节, 现场的同学踊跃提问, 就模型中提到的两次转移过程、显示表达的逼近等细节问题与 Yacine 教授进行了交流。Yacine 教授讲授过程深入浅出, 循循善诱, 与同学们进行了深入的探讨, 现场同学纷纷表示这次与 Yacine 教授的交流获益良多。

3. 北京师范大学陈木法院士讲座成功举办



3月21日, 应北京大学统计科学中心邀请, 中科院院士、北京师范大学陈木法教授在理科一号楼1114教室举办了题为“The charming maximal eigenpair”的学术讲座。陈木法教授于2003年和2009年当选为中国科学院和第三世界科学院院士, 2012年当选美国数学会 fellow。主要从事概率论与相关领域的研究工作。他所领导的概率论

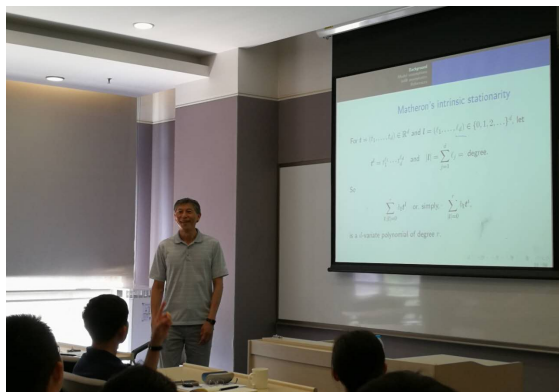
研究小组获得了2001~2010年共3期的国家自然科学基金委员会创新研究群体资助。两种主要数学评论杂志的评论一致认为他创建或领导了我国的一个“学派(School)”。获何梁何利基金“科学与技术进步奖”(2009)等12项奖励。

下午3时, 理科一号楼1114报告厅中, 北大师生济济一堂。讲座在陈松蹊教授的主持下拉开帷幕。陈木法院士从三个方面介绍了他最近关于最大特征对计算的研究进展。首先, 陈院士回顾了计算矩阵最大特征对的研究过程中最新获得的一些重要成果, 这些结果主要围绕已知算法中有效初值的选择展开, 而这些初值的选取思想又依赖于早前对 leading eigenpair 的估计结果。接下来, 陈院士向大家简要介绍了在不同情形下对 leading eigenpair 的基本估计方法, 尤其是一维情形下二阶椭圆型算子有关性质。最后, 陈院士提出了自己基于高效初值选择方法的迭代算法求特征对的算法, 在这一算法下, 原本复杂的求特征对问题只需要两次迭代便可以实现稳定结果。

讲座期间, 陈院士就初值选择的效率问题与该算法概率应用领域的细节问题与参加讲座的师生们进行了交流探讨, 并一一回答了师生们的问题。陈院士的讲座给大家带来的不仅是前沿的学术知识, 更引发了同学们对数学问题探索和钻研的兴趣。

4. 密歇根大学 Tailen Hsing 教授讲座成功举办

5月18日,应北京大学统计科学中心邀请,密歇根大学 Tailen Hsing 教授在北京大学光华管理学院二号楼 217 教室举办了主题为“Analyzing spatial data locally”的学术讲座。Tailen Hsing 教授 1984 年在北卡罗来纳州立大学获得博士学位,现任统计学顶级期刊《Annals of Statistics》联席主编,Hsing 教授的研究领域主要包括极值理论、相关数据的极限理论、函数型数据和空间数据分析。



下午 2 时,讲座在张俊妮老师的主持下开始。Tailen Hsing 教授首先介绍了平稳性假设在空间数据研究中的重要作用,他指出,如果数据是在一个充分小的区域中获取,平稳性假设将是一个较为合理的近似。因此,在处理非平稳空间数据时,将数据局部化进行分析可以将问题简化到较为熟悉的领域。随后,Hsing 教授介绍了目前较为成熟的平稳空间数据的建模方法,并详细介绍了将空间数据局部平稳的有关概念和理论结果。在

最后一部分,Hsing 教授介绍了更为一般化的非平稳空间模型,这一模型的协方差矩阵在局部上与 Matérn 协方差相似,针对这一模型,Hsing 教授给出了基于稠密数据进行模型推断的有关方法。

这次讲座共吸引了三十余名北大师参加,在提问环节,陈松蹊老师、黄辉老师、朱雪宁博士等多位师生先后提问,问题围绕研究过程中大家关心的局部平稳化对区域划分是否固定、局部平稳化的模型检验问题展开,现场讨论气氛热烈,大家纷纷表示对于空间统计领域的研究有了更加深刻的认识。

5. 曼彻斯特大学潘建新教授讲座成功举办



9月12日,应北京大学统计科学中心邀请,国家“国家级重大人才计划”入选者、曼彻斯特大学数学学院潘建新教授在理科一号楼1114教室举办了主题为“Joint mean-covariance modelling and its R package: jmcm”的学术讲座。潘建新教授现任英国皇家统计学会 Fellow,是世界著名的统计学杂志Biometrics和 Biometrical Journal 的Associated Editor,同时也是

International Statistical Institute 和 Institute of Mathematical Statistics 等国际学术组织的会员。潘建新教授曾担任曼大数学学院概率统计系主任,致力于统计学领域内复杂数据模型的理论研究及其在生物医学、金融及工业上的应用研究,取得了多项创新性研究成果。

下午 2 时,讲座在席瑞斌老师的主持下开始。潘建新教授首先强调,均值估计是一个比较传统的统计问题,而对于每个个体的协方差(方差)的估计问题某种程度上人们不够重视,例如英国某心脏病专家发在知名医学期刊《柳叶刀》上的一项研究表明,高血压虽对心脏病

有影响，但血压变化不稳定的患者更容易心脏出问题，所以此时研究方差比研究均值显得更有意义。针对每个样本的协方差估计问题，潘建新教授介绍了三种方法：modified Cholesky decomposition、alternative Cholesky decomposition、hyperpherical parameterization of Cholesky factor，这三种方法均能对协方差阵给出很好的估计，并且这三种方法的前两种经过 Cholesky 分解后的元素都有较好的统计解释性。例如 MCD 中得到的元素可解释为自回归系数和 innovation variance，ACD 中得到的元素可解释为滑动平均，从而可根据样本对这些元素进行估计从而得到正定的协方差矩阵。最后，潘老师现场演示了基于他的工作做出的 R package。

提问环节中，潘教授就三种方法效果的比较问题、数据的 order 问题与参加讲座的师生进行了交流探讨，并一一回答了师生们的问题。本次讲座吸引了众多北大师生参与，潘老师在传统统计问题上提出的新见解也给大家更多的启发。

6. 罗格斯大学陈嵘教授讲座成功举办

9 月 21 日，应北京大学统计科学中心邀请，罗格斯大学统计与生物统计系陈嵘教授在光华 2 号楼 217 教室举办了题为“Factor model for high dimensional matrix valued time series, with possible extension to tensor time series”的学术讲座。陈嵘教授是数理统计学会 (IMS) 资深会员，曾任美国伊州大学芝加哥分校商学院教授、北京大学光华管理学院商务统计与经济计量系教授系主任、美国国家自然科学基金委数学科学部项目主任。陈嵘教授于 20 世纪 90 年代初率先将非参数方法引入时间序列分析，并成为该领域权威；并于 90 年代后期提出了序贯蒙特卡罗方法的整体框架，奠定了序贯蒙特卡罗方法研究、拓展及广泛应用的基础，并成为该领域研究及应用的带头人。他在经济数据分析、旅游前景预测、股市分析、环境污染、电力系统分析各领域做出了有开创性的贡献。

下午 3 时 20 分，讲座在陈松蹊老师的主持下开始。陈嵘教授首先强调了高维矩阵时间序列数据的重要性，在实际的时间序列数据中，我们往往都能在一个时刻观察到一个矩阵，例如公司每季度的财务报表、每年多个国家的进出口贸易记录表等。在传统的统计方法中，我们通常是将一个矩阵展开获得一个维度较高的向量，但这样的做法往往会损失掉原本矩阵中结构信息，同时也会增加参数的维数从而增加估计难度。陈教授从因子模型入手，提出了一种新的矩阵因子模型，在公共因子矩阵 F 前后分别乘 A 、 B 两个载荷矩阵。这一做法既解决了原矩阵行与列之间的交互作用，又可以看做是一种分层模型的情形。在估计方法上，陈教授沿用了姚琦伟教授的思路，抓住误差项是白噪声的性质，最小化协方差阵的平方进行求解。随后，陈嵘教授又介绍了这一方法的极限性质与具体的案例应用。在陈教授介绍的 Fama-French 模型、公司财务数据与国家进出口数据中，基于矩阵因子模型取得的结果均比原本矩阵直接向量化得到的结果更加可靠。

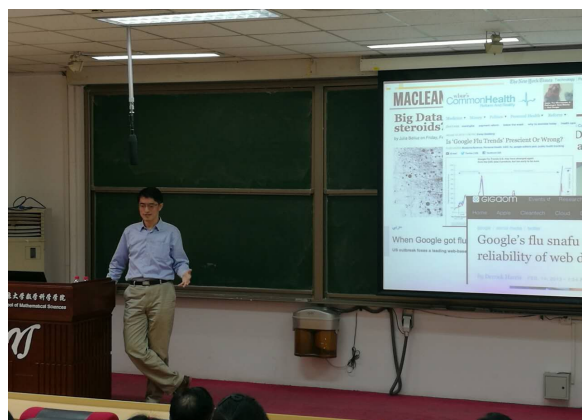
在提问环节，陈嵘教授就载荷阵的旋转、动态社交网络的应用困难、该方法在张量上的应用等问题与参加讲座的师生们进行了交流探讨，并一一回答了师生们的问题。本次讲座吸引了众多北大师生参与，大家在与陈老师的交流中对张量时间序列等研究热点问题有了更深入的了解。

Center for Statistical Science, Peking University

7. 哈佛大学 Samuel Kou 教授讲座成功举办

12月21日,应北京大学统计科学中心邀请,哈佛大学统计系 Samuel Kou 教授在理科一号楼 1114 教室举办了主题为“Big data, Google and disease detection: the statistical story”的学术讲座。Samuel Kou 是哈佛大学统计系教授、新加坡国立大学特聘访问教授,2007 年当选美国统计学会会士,2011 年当选国际统计学会会员,2012 年获统计学界的最高奖“考普斯会长奖”,他的主要研究方向包括大数据分析、非参数方法、生物化学中的随机推断、模型选择与经验贝叶斯方法等领域。

下午 2 时,讲座在姚方教授的主持下开始。Samuel Kou 教授从“大数据”的概念讲起,他指出,互联网中的海量数据不仅为学术界提供了丰富的研究课题,更是引起了工业界与政府组织的兴趣,这些互联网上的大数据给我们跟踪与预测大规模的社会活动提供了便利和可能性。Samuel Kou 教授向大家分享了流行病的跟踪研究。首先,他谈到了谷歌公司开发的“谷歌流感趋势”(Google Flu Trends)技术,这是一个基于大数据搜索进行流感预测的统计应用案例,在这一思路的基础上, Samuel Kou 教授提出了一个利用互联网在线信息估计流感爆发情况的新方法,这一方法在当前所有流感实时追踪模型中具有最突出的表现。利用这个方法, Samuel Kou 教授对亚洲与南美洲国家登革热疫情的精确追踪进行了应用,取得了理想的效果。



在提问环节, Samuel Kou 教授就变量选择、时间序列数据平稳性等大数据应用中出现的问题与参加讲座的师生们进行了交流探讨,并一一回答了大家的问题。本次讲座吸引了来自校内外近百名师生参加,大家对大数据时代统计学方法的发展与应用交换思路,发表看法,获得了进步。

8. 统计中心 2017 年来访人员名单

2017 年统计科学中心共接待海内外访问学者 70 次,参与学术报告,短期课程以及合作项目等各类学术活动。

来访人员	时间
Xiaohu Wang, The Chinese University of Hong Kong	2017-02-22~24
Yacine Ait-Sahalia, Princeton University	2017-03-10
Mu-Fa Chen Beijing Normal University	2017-03-21
Zhao Chunsheng, School of Physics, Peking University	2017-03-30
Xuewei Yang, Nanjing University	2017-04-06
簡錦漢, 台湾中央研究院, 经济研究所	2017-04-10
陳宜廷, 台湾中央研究院, 经济研究所	2017-04-10

張靜貞, 台湾中央研究院, 经济研究所	2017-04-10
許育進, 台湾中央研究院, 经济研究所	2017-04-10
周雨田, 台湾中央研究院, 经济研究所	2017-04-10
蕭代基, 台湾中央研究院, 经济研究所	2017-04-10
王成, 上海交通大学	2017-04-13
张忠元, 中央财经大学	2017-04-20
Dacheng Xiu, The University of Chicago	2017-04-26
凌少平, 志诺维思	2017-04-27
Edward I George, University of Pennsylvania	2017-05-10
Dong Li, Center for Statistical Science, Tsinghua University	2017-05-11
Mingxiu Hu, Takeda Pharmaceuticals and Yale University	2017-05-11
Tianxi Cai, Harvard University	2017-05-12
T.Tony Cai, University of Pennsylvania	2017-05-12
Matt Wand, University of Technology Sydney	2017-05-12
郑术蓉, 东北师范大学	2017-05-13
Dan Nettleton, Iowa State University	2017-05-13~17
Xinghua Zheng, Hong Kong University of Science and Technology	2017-05-13
Guangming Pan, Nanyang Technological University	2017-05-13
Cun-Hui Zhang, Rutgers University	2017-05-13
Regina Liu, Rutgers University	2017-05-13
Huibin Zhou, Yale University	2017-05-14
Heping Zhang, Yale University	2017-05-14
Ming Yuan, University of Wisconsin-Madison	2017-05-14
Ji Zhu, University of Michigan	2017-05-14
Pingshou Zhong, Michigan State University	2017-05-15
孙剑, 西安交通大学	2017-05-15
Yazhen Wang, University of Wisconsin-Madison	2017-05-16
Tailen Hsing, University of Michigan	2017-05-18
Xiaoou Li, University of Minnesota	2017-05-17~19
Chi Wang, University of Kentucky	2017-05-25
Ruiyan Luo, Georgia State University	2017-05-24~25
Yong Bao, Purdue University and University of Macau	2017-05-25

Balgobin Nandram, Worcester Polytechnic Institute, USA	2017-05-28~29
Frederick K. H. Phoa, Institute of Statistical Science, Academia Sinica	2017-06-01
Lingjiong Zhu, Florida State University	2017-06-08
Will Wei Sun, University of Miami	2017-06-13
Sijian Wang, University of Wisconsin-Madison	2017-06-13
Xueyan Zhao, Monash University	2017-06-21
Donald J. Wuebbles, University of Illinois Urbana-Champaign	2017-06-22
Jing Lei, Carnegie Mellon University	2017-06-27
Linbo Wang, Harvard University	2017-07-05/12-18
Weijie Su, University of Pennsylvania	2017-07-10
John Brant, 新西兰统计局	2017-08-20~09-30
Jianxin Pan, University of Manchester	2017-09-12
Tae-hwy Lee, University of California Riverside	2017-09-14
赵永鹏, 挖财信贷风控负责人	2017-09-21
Rong Chen, Rutgers University	2017-09-21
嵇少林, 山东大学	2017-10-12
Julie L. Forman, University of Copenhagen	2017-10-16
Bin Wang, Harbin Institute of Technology, Shenzhen	2017-10-26
贾慧珏, 深圳华大生命科学研究院宏基因组研究所	2017-11-15~17
Peter Song, University of Michigan	2017-11-27
Peng Ding, UC Berkeley	2017-12-07
Zhenhua Lin, University of California	2017-12-13
潘小川教授, 北京大学医学部	2017-12-14
Andreas Kryger Jensen, University of Copenhagen	2017-12-14
Xiucan Ding, University of Toronto	2017-12-19
Yiying Zhang, University of Hong Kong	2017-12-20
Samuel Kou, Harvard University	2017-12-21
Jiashun Jin, Carnegie Mellon University	2017-12-16~27
Zhide Fang, Louisiana State University	2017-12-26
Xinyi Xu, Ohio State University	2017-12-27
王少鹏, 北京大学	2017-12-28

七、论文发表

自 2017 年 1 月至 12 月, 统计科学中心教员共发表论文 42 篇, 按第一作者姓氏字母排序排列:

[1] Chang, J., Tan, W., Ling, Z., **Xi, R.**, Shao, M. ... Mao, W., Wu, C. and Lin, D. (2017). Comprehensive analysis of esophageal squamous-cell carcinoma identifies alcohol drinking-related mutation signature and novel genomic alterations, *Nature Communications*, 8:15290.

[2] Chen Y.(陈勇), **Liu Y.** (刘勇) (2017). On the fourth moment theorem for complex multiple Wiener–Ito integrals, *Infinite Dimensional Analysis, Quantum Probability and Related Topics* Vol. 20, No. 1 (2017) 1750005 (24 pages)

[3] Chen, Y., Jia, C. and **Jiang, D.Q.** (2017). Fluctuation-dissipation theorems for inhomogeneous Markov jump processes and a biochemical application, *J. Math. Phys.* 58, 023302.

[4] Chen, Z.-Q. **Ren, Y.-X.** and Yang, T. (2017). Law of large numbers for branching symmetric Hunt processes with measure-valued branching rates, *J. Theoret. Probab.* 30(3)(2017): 898–931.

[5] Dai, X., Müller, H. G., and **Yao, F.** (2017). Optimal Bayes classifiers for functional data and density ratios, *Biometrika*, 104, 545-560.

[6] Gao, Q., Wang, Z. C., Duan, M., Lin, Y. H., Zhou, X. Y., Worthley, D. L., ... **Xi R.**, Shi YY, Fan J. (2017). Cell Culture System for Analysis of Genetic Heterogeneity Within Hepatocellular Carcinomas and Response to Pharmacologic Agents, *Gastroenterology*, 152 (1), 232-242.

[7] Ge, H., Jia, C. and **Jiang, D.Q.** (2017). Cycle symmetry, limit theorems, and fluctuation theorems for diffusion processes on the circle. *Stochastic Process. Appl.* 127(6), 1897—1925.

[8] Guo, Nan; Wang, Fang; **Yang, Jingping** (2017). Remarks on composite Bernstein copula and its application to credit risk analysis, *Insurance: Mathematics and Economics*, 77, 38-48.

[9] He Y, Jia J, **Geng Z.** (2017). Structural learning of causal networks, *Behaviormetrika*, 2017, 44(1):287-305.

[10] Hydbring P, Wang Y, Fassl A, Li X, Matia V, Otto T, Choi YJ, Sweeney KE, Suski JM, Yin H, Bogorad RL, Goel S, Yuzugullu H, Kauffman KJ, Yang J, Jin C, Li Y, Floris D, Swanson R, Ng K, Sicinska E, Anders L, Zhao JJ, Polyak K, Anderson DG, **Li C.**, Sicinski P.

(2017) . Cell-Cycle-Targeting MicroRNAs as Therapeutic Tools against Refractory Cancers. *Cancer Cell*. 31(4):576-590.e8.

[11] Jiang, Bochuan; **Ai, Mingyao** (2017). Construction of uniform U-designs, *Journal of Statistical Planning and Inference* 181, 1-10.

[12] **Jiang, D.Q.**, Wang, Y. and Zhou, D.(2017). Phenotypic equilibrium as probabilistic convergence in multi-phenotype cell population dynamics. *PLoS ONE* 12(2): e0170916.

[13] Kong, Xiangshun; **Ai, Mingyao**; Mukerjee, Rahul (2017). Central limit theorems for four new types of U-designs, *Statistics* 51(3), 655-667.

[14] Kong, Xiangshun; **Ai, Mingyao**; Tsui, Kwok Leung (2017). Design for sequential follow-up experiments in computer emulations, *Technometrics*, online Feb. 3, 2017. (TCH-16-069.R2)

[15] Kong, Xiangshun; **Ai, Mingyao**; Tsui, Kwok Leung (2017). Flexible sliced designs for computer experiments, *Annals of the Institute of Statistical Mathematics*, online Feb. 25, 2017.

[16] Li, Yanfang, and **Jia Jinzhu** (2017). L1 least squares for sparse high-dimensional LDA, *Electronic Journal of Statistics*, 11.1(2017):2499-2518.

[17] Liu, H., Chiang, J. T. J., Fehr, R., **Xu, M.**, & Wang, S. (2017). How do leaders react when treated unfairly? Leader narcissism and self-interested behavior in response to unfair treatment, *Journal of Applied Psychology*, 102(11), 1590.

[18] **Ren, Y.-X.**, Song, R. and Zhang, R. (2017). Central limit theorems for supercritical branching nonsymmetric Markov processes, *Ann. Probab.*, 45(1)(2017): 564–623.

[19] **Ren, Y.-X.**, Song, R. and Zhang, R. (2017). Functional central limit theorems for supercritical superprocesses, *Acta Applicandae Mathematicae* 147(1)(2017):137–175.

[20] **Tu Yundong** (2017). Efficient Nonparametric Estimation in Simultaneous Equation Models, *Communications in Statistics - Theory and Methods*, 46, 3411-3416, April 2017.

[21] **Tu Yundong** (2017). On Spurious Regressions with Partial Unit Root Processes, *Economics Letters*, 150, 142-145, January 2017.

[22] **Tu Yundong** and Yi Yanping (2017). Forecasting Cointegrated Nonstationary Time Series with Time-varying Variance, *Journal of Econometrics*, 196(1), 83-98, January 2017.

- [23] Wang Y, Fan C, Zheng Y, **Li C.** (2017). Dynamic chromatin accessibility modeled by Markov process of randomly-moving molecules in the 3D genome, *Nucleic Acids Research*, 45(10):e85.
- [24] Wang, Yaping; **Ai, Mingyao** (2017). Definitive Screening Designs with Extreme Numbers of Level Changes, *Statistics and Probability Letters* 131, 13-18.
- [25] Wu P, Li T, Li R, Jia L, Zhu P, Liu Y, Chen Q, Tang D, Yu Y, **Li C.** (2017). 3D genome of multiple myeloma reveals spatial genome disorganization associated with copy number variations, *Nature Communications*, 2017 Dec 5;8(1):1937.
- [26] X. Qu, Lee L.F. and **Yu J.** (2017). QML Estimation of Spatial Dynamic Panel Data Models with Endogenous Time Varying Spatial Weights Matrices, *Journal of Econometrics*, 197, 173-201.
- [27] Xia Y., Liu Y., **Deng M., Xi R.** (2017). SVmine improves structural variation detection by integrative mining of predictions from multiple algorithms, *Bioinformatics*, 1;33(21):3348-3354.
- [28] Xia, Y., Liu Y., **Deng, M.** and **Xi, R.** (2017). Pysim-sv: a package for simulating structural variation data with GC-biases, *BCM Bioinformatics*, 18(Suppl 3):53.
- [29] Xie, Jiehua; Lin, Feng; **Yang, Jingping** (2017). On a generalization of Archimedean copula family, *Statistics & Probability Letters*, 125, 121-129.
- [30] Yang, Y., Wang, F., **Zhang, J.**, Xu, J., Yu, P. S. (2017). A topic model for co-occurring normal documents and short texts, *World Wide Web-Internet and Web Information Systems*, Published Online: 23 June 2017.
- [31] Yu, Jun; **Ai, Mingyao**; Wang, Yaping (2017). Optimal designs for linear models with Fredholm-type errors, *Journal of Statistical Planning and Inference*, online Oct. 16, 2017.
- [32] Yuan H., Xi R., Chen C. , **Deng M.** (2017). Differential Network Analysis via Lasso Penalized D-Trace Loss, *Biometrika*, 104 (4): 755-770.
- [33] Zhang S, Guo B, Dong A, He J, Xu Z, **Chen SX.** (2017). Cautionary tales on air-quality improvement in Beijing, *Proc. R. Soc. A* 20170457.
- [34] Zheng, Wei; **Ai, Mingyao**; Li, Kang (2017). Identification of universally optimal circular designs for the interference model, *Annals of Statistics* 45(4), 1462-1487.

- [35]Zhou Jing, **Tu Yundong**, Chen Yuxin and **Wang Hansheng** (2017). Estimating Spatial Autocorrelation with Sampled Network Data, *Journal of Business and Economic Statistics*, 35, 130-138, January 2017.
- [36]Zhou, J., Huang, D., **Wang, H.** (2017). A Dynamic Logistic Regression for Network Link Prediction, *Science China: Mathematics*, 60(1), 165-176.
- [37]Zhou, J., Tu, Y., Chen, Y., and **Wang, H.** (2017). Estimating spatial autocorrelation with sampled network data, *Journal of Business and Economics Statistics*, 35(1), 130-138.
- [38]Zhu, X., Pan, R., Li, G., Liu, Y. and **Wang, H.** (2017). Network Vector Autoregression, *The Annals of Statistics*, 45(3), 1096-1123.
- [39]Zou Tao & **Chen Song Xi** (2017). Enhancing Estimation for Interest Rate Diffusion Models With Bond Prices, *Journal of Business & Economic Statistics*, 35:3, 486-498
- [40]Zou, T., Lan, W., **Wang, H.**, and Tsai, C. L. (2017). Covariance Regression Analysis, *Journal of the American Statistical Association*, 112, 266-281.
- [41]王健, **张俊妮** (2017). 统计模型在中文文本挖掘中的应用,《数理统计与管理》,36(4)。
- [42]周静, 周小宇, **王汉生** (2017) . “自我网络特征对电信客户流失的影响研究”,《管理科学》,第 30 卷第 5 期, 28-37。



八、研究报告

- [1] Susanne Strohmaier, Nicolai Haase, Jørn Wetterslev, and Theis Lange
A simple to implement algorithm for natural direct and indirect effects in survival studies with a repeatedly measured mediator
- [2] Yumou Qiu, **Song Xi Chen** and Dan Nettleton
Detecting rare and faint signals via thresholding maximum likelihood estimators
- [3] **Wang Miao** and Eric Tchetgen Tchetgen
Invited Commentary: Bias Attenuation and Identification of Causal Effects With Multiple Negative Controls
- [4] **Wang Miao** and Eric Tchetgen Tchetgen
Identification and Inference With Nonignorable Missing Covariate Data
- [5] **Wang Miao**
Identifying Causal Effects With Proxy Variables of an Unmeasured Confounder
- [6] XB Ding and **XH Zhou**
Nonparametric transformation regression with unknown variance function for highly skewed data
- [7] C. Zheng and **XH Zhou**
Causal mediation analysis on failure time outcome without sequential ignorability.
- [8] Y Wu, Y Li, Y Hou, K Li, and **XH Zhou**
Study duration for three-arm non-inferiority survival trials designed for accrual by cohorts. Statistical Methods in Medical Research
- [9] Jiang Chang, Wenle Tan, Zhiqiang Ling, **Ruibin Xi**, Mingming Shao, Mengjie Chen, Yingying Luo, Yanjie Zhao, Yun Liu, Xiancong Huang, Yuchao Xia, Jinglin Hu, Joel S Parker, David Marron, Qionghua Cui, Linna Peng, Jiahui Chu, Hongmin Li, Zhongli Du, Yaling Han, Wen Tan, Zhihua Liu, Qimin Zhan, Yun Li
Comprehensive analysis of esophageal squamous-cell carcinoma identifies alcohol drinking-related mutation signature and novel genomic alterations
- [10] Yuchao Xia, Yun Liu, Minghua Deng and **Ruibin Xi**
Pysim-sv: a package for simulating structural variation data with GC-biases
- [11] Haiyang Liu, Jack Ting-Ju Chiang, Ryan Fehr, **Minya Xu** and Siting Wang
How do leaders react when treated unfairly? Leader narcissism and self-interested behavior in response to unfair treatment.

- [12] X. Dai, G.H. Müller and **F. Yao**
Optimal Bayes classifiers for functional data and density ratios
- [13] Z. Lin, G.H. Müller, and **F. Yao**
Mixture inner product spaces and their application to functional data analysis
- [14] Yuanpei Cao, **Wei Lin** and Hongzhe Li
Two-sample tests of high-dimensional means for compositional data
- [15] **Yanxia Ren**, Renming Song and Zhenyao Sun
A 2-spine decomposition of the critical Galton-Watson tree and a probabilistic proof of Yaglom's theorem
- [16] A.E. Kyprianou , Sandra Palau and **Yanxia Ren**
Almost sure growth of supercritical multi-type continuous state branching process
- [17] **Yanxia Ren**, Renming Song and Rui Zhang
Supercritical superprocesses: proper normalization and non-degenerate strong limit
- [18] Zhen-Qing Chen, **Yanxia Ren** and Renming Song
Llog L criterion for a class of multitype superdiffusions with non-local branching mechanisms
- [19] Zhen-Qing Chen, **Yanxia Ren** and Ting Yang
Skeleton decomposition and law of large numbers for supercritical superprocesses
- [20] Xiaojun Mao, **Song Xi Chen** and Raymond K. W. Wong
Matrix Completion with Covariate Information
- [21] M. Koudstaal, and **F. Yao**
From Multiple Gaussian Sequences to Functional Data and Beyond: A Stein Estimation Approach
- [22] M. Koudstaal, and **F. Yao**
Supplementary Material for “From Multiple Gaussian Sequences to Functional Data and Beyond: A Stein Estimation Approach”
- [23] Shuyi Zhang, Bin Guo, Anlan Dong, Jing He, Ziping Xu, **Song Xi Chen**
Cautionary Tales on Air Quality Improvement in Beijing
- [24] Chong Chen, **Ruibin Xi** and Nan Lin
Community Detection by L0-penalized Graph Laplacian

- [25] Wei Lan, Yingying Ma, Junlong Zhao, **Hansheng Wang** and Chih-Ling Tsai
Sequential Model Averaging for High Dimensional Linear Regression Models
- [26] Lujun Li, Hui Shao, Ruodu Wang and **Jingping Yang**
Worst-case Range Value-at-Risk with Partial Information
- [27] D. Huang, X. Chang and **Hansheng Wang**
Spatial Autoregression with Repeated Measurements for Social Networks
- [28] Yacine Ait-Sahalia, **Chenxu Li**
Closed-Form Implied Volatility Surfaces for Stochastic Volatility Models
- [29] Shuyi Zhang, Bin Guo, Hengfang Wang, Jing He, Anlan Dong, Ziping Xu, **Wei Lin**,
Songxi Chen
空气质量评估报告(三)
- [30] Lei Chen, Bin Guo, Jiasheng Huang, Shuyi Zhang, Hengfang Wang, Jing He, **Songxi Chen**
空气质量评估报告(四)
- [31] Huiming Zhang, Bo Li and G. Jay Kerns
A characterization of signed discrete infinitely divisible distributions
- [32] Tianyuan Zhou and **Jinzhu Jia**
Basis-expansion Random-Projection Ensemble for Regression
- [33] Yuling Yan and **Jinzhu Jia**
Sparse logistic regression with penalized weighted score function
- [34] **Yundong Tu**, Qiwei Yao and Rongmao Zhang
Error-Correction Factor Models for High-dimensional Cointegrated Time Series
- [35] **Yundong Tu** and Ying Wang
Adaptive Estimation of Functional-coefficient Cointegration Models with Nonstationary Volatility
- [36] **Yundong Tu**
On spurious regressions with partial unit root processes

九、学术讲座

北京大学统计科学中心 2017 年共组织学术报告 45 场，报告人来自美国、英国、丹麦、加拿大、香港以及国内各高校、研究机构和知名企业。其中来自海外的学者 30 场，国内的学者 15 场。学术报告拓宽了老师和学生的学术视野，促进了统计学领域的学术交流，活跃了统计中心的学术氛围。

[1] New Distribution Theory for the Estimation of Structural Break Point in Mean

Xiaohu Wang, The Chinese University of Hong Kong

2017-02-23 14:00-15:30

光华 2 号楼 217 教室

[2] An Introduction to Expansion Methods for the Transition Density of Continuous-Time Models

Yacine Ait-Sahalia, Princeton University

2017-03-10 14:00-15:30

光华 1 号楼 112 教室

[3] The charming maximal eigenpair

Mu-Fa Chen, Beijing Normal University

2017-03-21 15:00-16:00

理科一号楼 1114 教室

[4] What have we learned from HaChi (HAZE IN CHINA) project?

Zhao Chunsheng, Peking University

2017-03-30 14:00-15:00

光华 2 号楼 217 教室

[5] Can Financial Innovation Succeed by Catering to Behavioral Preferences? Evidence from a Callable Options Market

Xuewei Yang, Nanjing University

2017-04-06 14:00-15:00

光华 2 号楼 217 教室

[6] Random Matrix Theory and Large Sample Covariance Matrix

王成, 上海交通大学

2017-04-13 14:00-15:00

理科一号楼 1114 教室

[7] On Equivalence of Likelihood Maximization of Stochastic Block Model and Nonnegative Matrix Factorization, and Beyond

张忠元, 中央财经大学

2017-04-20 14:00-15:00

理科一号楼 1114 教室

[8] Taming the Factor Zoo

Dacheng Xiu, The University of Chicago

2017-04-26 14:00-15:00

光华酒店 K01

[9] Identifying Tissue Origin of Cancer Cells with Somatic Mutations and Copy Number Alterations

凌少平, 志诺维思

2017-04-27 14:00-15:00

理科一号楼 1303 教室

[10] Inference in Threshold GARCH Models

Dong Li, Tsinghua University

2017-05-11 14:00-15:00

光华 2 号楼 217 教室

[11] Statistical Issues in Drug Development

Mingxiu Hu, Takeda Pharmaceuticals and Yale University

2017-05-11 15:30-16:30

理科一号楼 1365 教室

[12] Analyzing spatial data locally

Tailen Hsing, University of Michigan

2017-05-18 14:00-15:00

光华 2 号楼 217 教室

[13] Generalized Likelihood Ratio Test of Separate Families

Xiaou Li, University of Minnesota

2017-05-19 14:00-15:00

理科一号楼 1418 教室

[14] Causal Effect Estimation While Accounting for Uncertainty in Confounder and Effect Modifier Selection

Chi Wang, University of Kentucky

2017-05-25 10:00-11:00

理科一号楼 1114 教室

[15] Function-on-Function Regression with Thousands of Predictive Curves

Ruiyan Luo, Georgia State University

2017-05-25 14:00-15:00

理科一号楼 1114 教室

[16] Indirect Inference Estimation of Dynamic Panel Data Models

Yong Bao, Purdue University, USA and University of Macau, China

2017-05-25 15:30-16:30

光华 2 号楼 217 教室

[17] Bayesian Predictive Inference for Numerous Sub-Areas Using Logistic Regression

Balgobin Nandram, Worcester Polytechnic Institute, USA

2017-05-29 14:00-15:00

光华 2 号楼 216 教室

[18] Orthogonal Arrays with Circulant Property: Construction, Analysis and Applications to fMRI Experiments

Frederick K. H. Phoa, Institute of Statistical Science, Academia Sinica

2017-06-01 14:00-15:00

理科一号楼 1114 教室

[19] Approximate Variational Estimation for a Model of Network Formation

Lingjiong Zhu, Florida State University

2017-06-08 14:00-15:00

光华 2 号楼 217 教室

[20] Provable Sparse Tensor Decomposition for Personalized Recommendation and Dynamic Clustering

Will Wei Sun, University of Miami

2017-06-13 14:00-15:00

光华 2 号楼 216 教室

[21] Sparse Additive Index Model for High-Dimensional Data with Grouped Covariates

Sijian Wang, University of Wisconsin-Madison

2017-06-13 16:00-17:00

理科一号楼 1418 教室

[22] A Framework for Distribution-Free Regression

Jing Lei, Carnegie Mellon University

2017-06-27 14:00-15:00

理科一号楼 1114 教室

[23] Variation Independent Parameterizations of Binary Causal Models

Linbo Wang, Harvard University

2017-07-05 14:00-15:00

理科一号楼 1114 教室

[24] Multiple Testing and Adaptive Estimation via the Sorted L-One Norm

Weijie Su, University of Pennsylvania

2017-07-10 14:00-15:00

理科一号楼 1114 教室

[25] Joint mean-covariance modelling and its R package: jmcm

Jianxin Pan, University of Manchester

2017-09-12 14:00-15:00

理科一号楼 1114 教室

[26] "Double Boosting for High Dimensional IV Regression Models" (coauthored with Hao Xu) and "Asymmetric AdaBoost for High Dimensional Maximum Score Regression" (coauthored with Jianghao Chu and Aman Ullah)

Tae-hwy Lee, University of California Riverside

2017-09-14 14:00-15:30

光华 2 号楼 217 教室

[27] 基于客户生命周期价值(LTV)的风控策略选择-----一个增强学习的案例分享

赵永鹏, 挖财信贷风控负责人

2017-09-21 14:00-15:00

光华 2 号楼 217 教室

[28] Factor model for high dimensional matrix valued time series, with possible extension to tensor time series

Rong Chen, Rutgers University

2017-09-21 15:10-16:10

光华 2 号楼 217 教室

[29] The least squares estimator of random variables under sublinear expectations

嵇少林, 山东大学

2017-10-12 14:00-15:00

光华 2 号楼 217 教室

[30] Non-parametric bootstrapping of diffusion models with application to testing the Markov hypothesis

Julie L. Forman, University of Copenhagen

2017-10-16 14:00-15:00

光华 2 号楼 217 教室

[31] Uniform and L_p Convergences for Nonparametric Kernel Estimation of Diffusion Models

Bin Wang, Harbin Institute of Technology, Shenzhen

2017-10-26 14:00-15:00

光华 2 号楼 217 教室

[32] Making sense of the microbiome through numbers

贾慧珏, 深圳华大生命科学研究院宏基因组研究所

2017-11-16 14:00-15:00

理科一号楼 1418 教室

[33] Renewable estimation and incremental inference in generalized linear models for streaming datasets

Peter Song, University of Michigan

2017-11-27 14:00-15:00

理科一号楼 1114 教室

[34] Randomization Inference for Treatment Effect Heterogeneity

Peng Ding, UC Berkeley

2017-12-07 14:00-15:00

理科一号楼 1418 教室

[35] Intrinsic Riemannian Functional Data Analysis

Zhenhua Lin, University of California

2017-12-13 14:00-15:00

理科一号楼 1114 教室

[36] MAIAC-based Long-term Trends in Ambient PM_{2.5} and the Effects on Population Mortality in Beijing, China

潘小川教授, 北京大学医学部

2017-12-14 14:00-15:00

理科一号楼 1479 教室

[37] Non- and semi-parametric Object-oriented Data Analysis for random density functions

Andreas Kryger Jensen, University of Copenhagen

2017-12-14 15:15-16:15

光华 2 号楼 217 教室

[38] Causal Inference with Unmeasured Confounding: an Instrumental Variable Approach

Linbo Wang, Harvard University

2017-12-18 14:00-15:00

理科一号楼 1114 教室

[39] Random matrix theory: From mathematical physics to high dimensional matrix estimation and time series analysis

Xiukai Ding, University of Toronto

2017-12-19 15:00-16:00

理科一号楼 1493 教室

[40] Risk-adjusted Bowley Reinsurance under Distorted Probabilities

Yiying Zhang, University of Hong Kong

2017-12-20 9:00-11:00

理科一号楼 1114 教室

[41] Big data, Google and disease detection: the statistical story

Samuel Kou, Harvard University

2017-12-21 14:00-15:00

理科一号楼 1114 教室

[42] Coauthorship and Citation Networks of Statisticians

Jiashun Jin, Carnegie Mellon University

2017-12-26 14:00-15:00

光华 2 号楼 217 教室

[43] An effective algorithm for big streaming data

Zhide Fang, Louisiana State University

2017-12-26 15:10-16:10

光华 2 号楼 217 教室

[44] Estimating Heterogeneous Causal effects: A Bayesian Nonparametric Approach

Xinyi Xu, Ohio State University

2017-12-27 10:30-11:30

理科一号楼 1418 教室

[45] Mathematical and Statistical Approaches in Ecology

王少鹏, 北京大学

2017-12-28 14:00-15:00

理科一号楼 1114 教室



十、委员会



王永雄

主席，斯坦福大学统计系及健康研究和政策系教授，美国科学院院士，曾获得过统计学界最高奖“考普斯会长奖”。



范剑青

委员，普林斯顿大学运筹与金融工程系系主任，统计研究委员会主任，中央研究院院士，曾获得过统计学界最高奖“考普斯会长奖”、洪堡研究奖、世界华人数学家大会晨兴数学金奖等奖项。



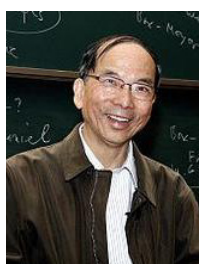
马志明

委员，中国科学院数学与系统科学研究院应用数学研究所研究员，国家“973”计划重点基础研究发展规划“核心数学的前沿问题”项目首席科学家，曾任中国数学会理事长，中国科学院院士，第三世界科学院院士。



Peter J. Bickel

委员，加州大学伯克利分校统计系教授，美国艺术与科学院院士，美国科学院院士。



吴建福

委员，乔治亚理工学院工程统计讲座教授，数理统计学会、美国统计学会、运筹学和管理学研究协会等学会会士，美国工程院院士，曾获得过统计学界最高奖“考普斯会长奖”。



黎子良

委员，现为美国斯坦福大学教授。1983 年获国际统计学界的“考普斯会长奖”。

2. 科学委员会



陈嵘

委员，罗格斯大学，北京大学光华管理学院特聘教授，美国统计学会荣誉会员，美国统计学会学报、商务经济统计学报编委，兼泛华统计学会董事。



陈松蹊

委员，北京大学统计科学中心联席主任，北京大学光华管理学院商务统计与经济计量系联席系主任，国家首批“国家级重大人才计划”入选者，现为北京大学讲席教授。



刁锦寰

委员，国际著名统计学家，现为美国芝加哥大学商学院经济计量学与统计学荣休讲席教授。他是泛华统计协会的创始人暨首任会长，是美国统计学会高级会员、台湾中央研究院院士，北京大学名誉教授。



耿直

委员，北京大学统计科学中心联席主任，北京大学数学科学学院概率统计系教授，获评国家级重大人才计划（基金委）。现任中国现场统计研究会第九界理事会理事长（2013-2017），IMS-China 主席，中国统计学会副会长，国家社会科学基金学科评审组专家等职。



胡永华

委员，北京大学公共卫生学院常务副院长、流行病与卫生统计学系主任，教育部流行病学重点实验室主任，北京大学循证医学中心常务副主任，北京大学公共卫生与预防医学教育委员会主任委员，被聘为香港中文大学荣誉教授，亚太地区公共卫生科学理事会中国区域主任。



陈大岳

委员，北京大学数学科学学院院长，中国数学会秘书长，中国概率统计学会副理事长。曾获国家级重大人才计划（基金委）等多项资助，曾任中国数学会副理事长和《数学进展》主编等职。



姚琦伟

委员，伦敦政治经济学院统计系主任，北京大学光华管理学院特聘教授，他还是皇家统计学会荣誉会员、国际统计研究所推选成员、美国统计协会、数理统计学会、泛华统计学会会员。

3. 指导委员会



高松

主席，北京大学党委常委、副校长，中国科学院化学学部院士，英国皇家化学学会会士(FRSC，2007-)，2013年当选为发展中国家科学院院士，2013年荣获“何梁何利基金科学与技术进步奖”。



陈大岳

委员，北京大学数学科学学院院长，中国数学会秘书长，中国概率统计学会副理事长。曾获国家级重大人才计划（基金委）等多项资助，曾任中国数学会副理事长和《数学进展》主编等职。



刘俏

委员，北京大学光华管理学院院长，国家级重大人才计划2项（教育部、基金委）人选。兼任深圳证券交易所专家评审委员会委员和中国证监会、中国金融期货交易所、民生银行、及深交所博士后站指导导师。在公司金融，实证资产定价、实际期权、市场微观结构和中国经济研究等方面拥有众多著述。



张平文

委员，现任北京大学副教务长、学科建设办公室主任，数学科学学院教授、博士生导师，中国科学院院士，北京大学数学及其应用教育部重点实验室主任。曾获国家自然科学二等奖、教育部高等学校自然科学一等奖、国家级重大人才计划2项（教育部、基金委）人选、冯康科学计算奖、国家自然科学基金委“创新研究群体”学术带头人等多项荣誉。



Tel: 0086-010-62760736
Fax: 0086-010-62760736
Mail: stat-center@pku.edu.cn
Http: www.stat-center.pku.cn
Add: 北京大学理科五号楼
Postcode: 100871