

# 北京大学统计科学中心

Center for Statistical Science, Peking University

2015年度报告

祝贺北大统计科学中心的成立，  
希望“中心”成为吸引海外高层次人才，  
培养本土拔尖人才的基地。  
李源澄  
6.18

目 录

一、综述 .....	2
二、人员组成 .....	3
三、中心要闻 .....	12
四、交流访问 .....	27
五、发表论文 .....	32
六、研究报告 .....	36
七、学术讲座 .....	39
八、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、光华管理学院以及医学部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。

中心的成立得到了各级政府部门和北京大学相关部门的支持和指导，他们在中心向着建设具有国际先进水平的队伍、培养复合型统计人才、大力推进跨学科研究、促进高校与政府及企业等机构的积极互动的目标不断迈进的过程中也一直给予关注与支持。

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，整合全校的统计研究力量，致力于培养综合型的统计人才，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、世界一流的统计研究机构。

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；执行委员会负责中心的日常管理工作。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”。

统计科学中心的主要任务是建立一支具有国际先进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验设计和统计分析水平，为这些学科的发展提供数据收集和数据分析的统计方法。

中心的成立将在北京大学建立一个吸引海内外优秀统计人才的基地。中心非常重视开放式、国际化的学术环境建设，广泛联系全球各地的统计学者，吸引优秀人才到北京大学工作。目前，统计科学中心正在积极开展研究生招生工作，几年之后，中心将为我国的统计学、自然科学、人文社会科学等各个领域提供杰出的科学研究人才，为各行业提供优秀的管理和技术人才，进一步提高北京大学在国内统计学研究领域的领先优势。



Peter G. hall 教授摄

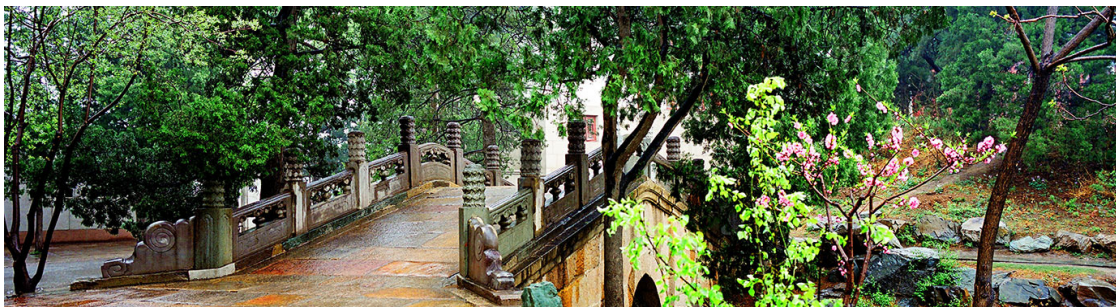


## 一、年度综述

2015 年统计科学中心继续在统计学研究、博士生培养、及推动北大及中国统计学发展等方面开展工作。

统计科学中心积极开展同国际和国内统计学界的学术交流与合作。1 月份，举办了“数量金融国际会议”和“首届环亚太青年计量学者 (YEAP) 会议”。5 月份，北京大学名誉教授刁锦寰教授开设以“向量 ARIMA 模型和多元时间序列的实例应用”为主题的统计学课程，国际著名统计学家 Alan E Gelfand 教授开设以“贝叶斯推断”为主题的统计学课程，本次短期课程的内容既包含贝叶斯推断的基础内容，也包含了贝叶斯统计的前沿研究问题。为促进海峡两岸学者的交流合作，陈松蹊教授和涂云东博士筹办和组织了第二届海峡两岸经济计量和环境统计交流与对话。6 月 6 日-7 日，王汉生教授筹办和组织了第八届中国 R 语言会议暨 2015 北大光华数据与价值论坛成功举办，6 月 29-30 日宋晓军博士筹备组织了“计量经济学中的模型设定检验”短期课程，6 月 25-28 日“国家级重大人才计划”学者蔡天文教授和林伟博士筹办和组织了“大数据时代的高维统计”国际青年学者会议暨夏季短期课程。7 月份，由陈大岳教授、任艳霞教授和刘勇教授筹办的“2015 年北京大学青年概率论坛”及短期课程成功举行。7 月统计中心成功举办了第二届优秀大学生夏令营。2015 年 9 月，中心首批两名博士研究生入学。统计中心资助和举办的这些活动极大地扩大了中心在国内外的影响力。

2015 年中心教员共发表 34 篇高质量学术论文，另有 14 篇已被接收待刊登。据最新统计北大统计学者在 2012-2015 四年间在统计学国际四大顶尖刊物上发表 19 篇论文（实为 20 篇），名列全国各高校之首，其中中心近年引进的老师们对此做出了有益的贡献。三月份由陈松蹊教授和黄辉博士带领的中心环境统计团队发布“北京地区 2010-2014 年 PM2.5 污染状况研究”报告，在社会上和政府大气环境管理领域有很大反响；十月份，基于该空气质量报告的文章被英国皇家学会会刊发表，中心环境统计团队成功地申请到国家重点研究专项：空气质量统计诊断模型，由陈松蹊教授带领的主要由统计中心教员组成的研究团队成功申请了国家自然科学基金重点项目“大数据驱动的管理决策模型与算法”。王汉生教授入选国家级重大人才计划（基金委）。他是北大继耿直，陈大岳之后概率统计学方向的第三位入选者。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，举办了学术报告共 49 场，其中来自海外的学者 26 场，国内的学者 13 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，对中心的学术发展起到了很大的推动作用。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、海内外同行。北京大学统计科学中心所取得的成绩都凝聚着众人的智慧和汗水，得益于各界的帮助和支持，我们将继续努力建设国内领先、世界一流的统计研究机构。





## 二、人员组成

### 1. 执行委员会

陈松蹊

统计科学中心联席主任, 光华管理学院商务统计与经济计量系联席系主任, 首批“国家级重大人才计划”入选者, 北京大学讲席教授。美国统计学会会士, 数理统计协会资深会员, 国际统计学会当选会员, 中国统计学会常务理事。研究方向: 数理统计, 经济计量学, 高维数据分析, 环境统计。

耿直

统计科学中心联席主任, 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)。曾任 IMS-China 主席, 中国现场统计研究会第九届理事会理事长(2013-2017), 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。研究方向: 因果推断, 数理统计, 生物医学统计。

陈大岳

统计科学中心副主任, 曾获国家级重大人才计划(基金委)等多项资助, 现任北京大学数学科学学院常务副院长、中国数学会副理事长和《数学进展》主编等职。研究方向: 马氏过程, 无穷粒子系统, 渗透模型, 随机游动。

房祥忠

统计科学中心执行委员会委员, 现任北京大学数学科学学院概率统计系系主任。研究方向为生存分析和可靠性。1996年获中国航天总公司(部级)科技进步二等奖。2002年获北京市科技进步二等奖。2002年获中国国防科学技术工业委员会国防科学技术奖三等奖。

李程

统计科学中心执行委员会委员。研究方向: 生物信息和计算生物学在癌症基因组学中的应用。研究组开发的 dChip、ComBat 软件和相关数据分析方法被广泛应用于基因表达和 SNP 生物芯片的数据分析和可视化。

王汉生

统计科学中心执行委员会委员, 光华管理学院商务统计与经济计量系联席系主任。美国统计学会会士, 现为国际统计研究所、美国数理统计研究所、英国皇家统计协会以及泛华统计学会会员。研究方向: 高维数据分析, 变量选择, 搜索引擎营销, 社会关系网络。

吴岚

统计科学中心执行委员会委员, 数学科学学院金融数学系系主任。研究方向: 金融资产定价, 金融统计, 经济资本模型, 精算学。中国精算师协会会员, 教育委员会委员, 中国精算师考试负责人。

## 2. 全职教员



**贾金柱研究员:** 2009 年在北京大学概率统计系获博士学位, 2009-2010 年在 UC Berkeley 统计系进行博士后研究工作, 2011 年加入北大统计中心和数学学院。贾金柱博士的研究方向为高维统计推断、统计机器学习和因果推断等。他在变量选择方法的理论研究、高维数据统计学习的应用和因果推断等领域发表论文多篇。

高维统计中的预处理方法:

Lasso 这一方法的大样本性质(相合性)需要条件。它的符号相合(模型选择相合性)需要 irrepresentable condition, 它的 L2 相合性需要限制特征值条件 (restricted eigenvalue condition)。Jinzhu Jia and Karl Rohe (2015)的工作 Preconditioning the Lasso for sign consistency. Electronic Journal of Statistics, pp 1150-1172. 展示和提出如何规避稀疏线性回归分析需要的条件。给定  $X, Y$  满足如下线性回归方程:

$$Y = X\beta + \epsilon$$

该研究提出当设计矩阵  $X$  不满足 irrepresentable condition 时, 通过选择一个合适的矩阵  $F$ , 对方程进行如下的预处理: 回归方程两边都乘以  $F$ :

$$FY = FX\beta + F\epsilon$$

去掉 Lasso 所需要的 irrepresentable condition 和 restricted eigenvalue condition. 该研究提出的  $F$  如下构成: 记  $X$  的 SVD 分解为

$$X = UDV^T \text{ 则 } F \text{ 定义为 } F = UD^{-1}U^T.$$

由于  $F$  的定义取决于  $X$  的 SVD 分解, 我们称这种预处理方法为基于 SVD 分解的预处理方法。我们的预处理使得变换后的设计矩阵( $FX$ ) 的特征值只有 0 或者 1。当变量的个数( $p$ )小于观测数据( $n$ )的个数时, 变换后的矩阵是正交矩阵, 这个时候 irrepresentable condition 和 restricted eigenvalue condition 都成立。当变量的个数大于观测数据个数的时候, 变换后的设计矩阵行正交, 此时我们也证明了 irrepresentable condition 和 restricted eigenvalue condition 都以很大的概率成立。我们证明了在高维回归中, 预处理可是规避 irrepresentable condition, 并且明显改进 Lasso 的性质。随机模拟也进一步证明了我们的理论。我们的早先 arxiv 版本是 Preconditioning to comply with the irrepresentable condition. google scholar 显示, 该结果目前已经被引用 11 次。





**涂云东研究员:** 2012 年在加州大学获经济学博士学位，同年加入北大统计中心和光华管理学院。涂云东博士的研究兴趣涵盖非参数建模，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。他的研究致力于如何利用已有的信息或数据进行更加有效的推断和经济预测，让数据自身来揭示背后的经济规律，减少经济研究中作出的不合理假设。

股权溢价的非线性预测模型和经济约束：

涂云东研究员和他的合作者 2015 年在 *Journal of Business and Economic Statistics* 上发表论文 (Lee, Tu and Ullah, 2015 JBES) 研究了金融学中资产溢价的非线性建模和预测。资产溢价的预测是实证金融研究中的热点问题，长期以来，大量的研究文献通过实证发现，资产溢价不可以被金融学中的其它变量所预测。该研究发现资产溢价和金融变量间存在非线性关系，而这一非线性关系在文献中并没有被发现。进一步，该研究结合实证金融中资产溢价非负性的约束条件，建立了局部平均的非线性预测模型以及带约束的局部平均非线性预测模型，有效的改善了文献中一直以来的最佳预测---历史平均预测。同时，该研究还利用自助平均 (bootstrap aggregating or bagging) 的方法对约束预测进行模型平均，有效的降低了约束预测模型的预测均方差。对所提出的新的预测方法的理论性质，该研究进行了渐近理论推导，推导了新的估计量的渐近偏差、方差和均方误差对局部参数的依赖关系，发现在大多数局部参数的取值空间上，新的估计量在均方误差上都有更好的表现。实证分析表明，在样本外的预测精度上，新的预测估计量相比历史平均预测有较大改善。

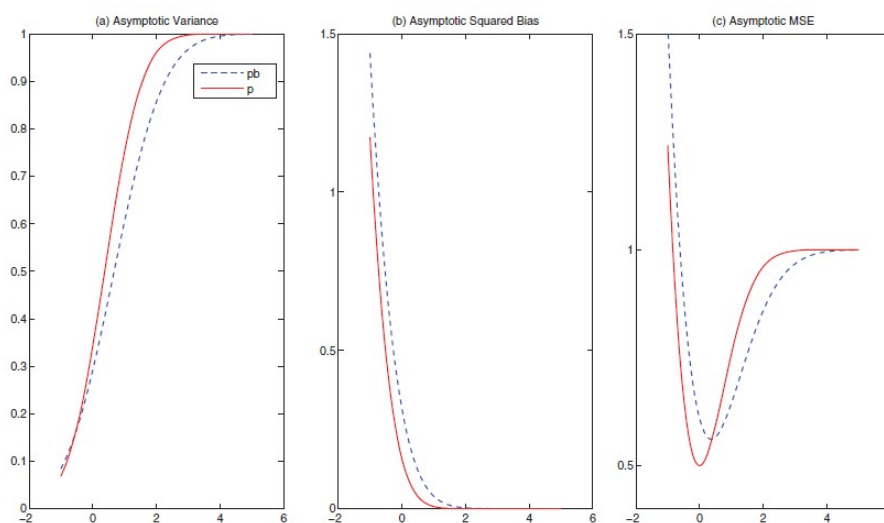
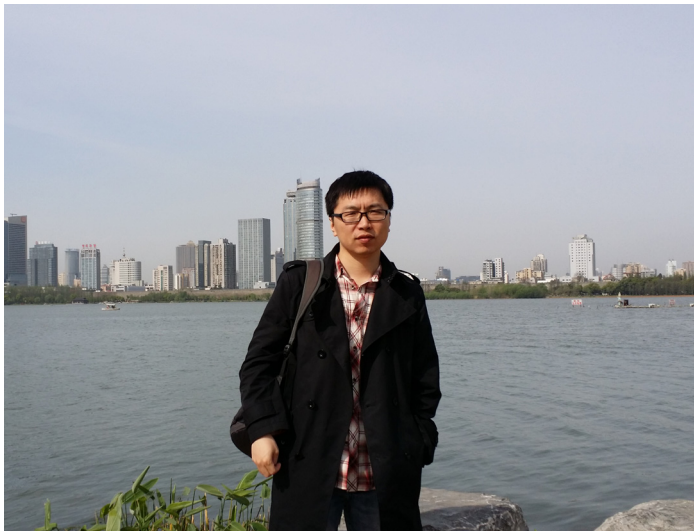


Figure 1. Asymptotic properties of the local historical average with positivity constraint ( $LHA_p$ , solid line) and the local historical average with positivity constraint and bagging ( $LHA_{pb}$ , dashed line).



### 席瑞斌研究员:

2009 年在华盛顿大学获数学博士学位，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。

### 多功能干细胞技术:

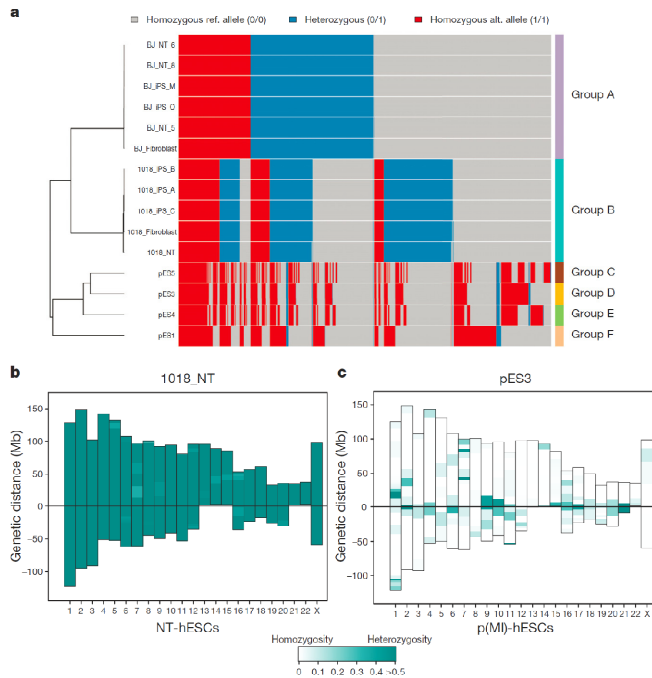


Figure 1 a. 通过分析诱导多功能细胞基因组及其来源体细胞基因组，可以排除其他细胞的污染；b,c 两种不同的 iPS 诱导技术所得到的多功能细胞具有不同的基因组特征。

席瑞斌研究员和他的合作者在 Nature 上发文 (De Los Angeles et al. 2015 Nature) 研究了多种多功能干细胞技术。多功能干细胞技术是干细胞领域的一项重大突破，在生命科学的基础研究中占有重要的地位，有广泛的医学应用前景。然而，由于 2014 年日本理化研究所的小保方晴子涉嫌造假的事件后，如何确定通过多功能干细胞技术得到的细胞为多功能细胞引起了相关研究人员的重视。本文回顾了

多种多功能干细胞的验证方法，并提出了可以通过分析高通量测序数据的方法来确定所得到的细胞无其他细胞污染 (上图 a)。同时，由于不同技术得到的多功能细胞的基因组会有不同的特点，我们也可以通过分析高通量测序数据来确定所得到细胞确实具有所该特点 (上图 b,c)。在图 a 中，我们可以清晰的看到由诱导多功能技术 (induced pluripotent stem cell 或 iPS 细胞) 及核转移技术 (nuclear transfer cell) 所生成的多功能细胞与其来源细胞有几乎一样的基因型 (如 BJ\_NT\_6 与其来源细胞 BJ\_Fibroblast, 1018\_iPS\_A 与其来源细胞 1018\_Fibroblas)，而由 parthenogenetic 技术产生的细胞 (如 pES1)，则由于 parthenogenetic 技术的特点，几次实验所产生多功能干细胞的基因型有较大不同。在图 c 中显示了由 parthenogenetic 技术产生的多功能干细胞的只在靠近 centromere 附件有异源 SNP，其余则大部分为同源 SNP，此特点与 parthenogenetic 技术点一直。相对的，由核转移技术生成的多功能干细胞则无此特点 (图 b)。

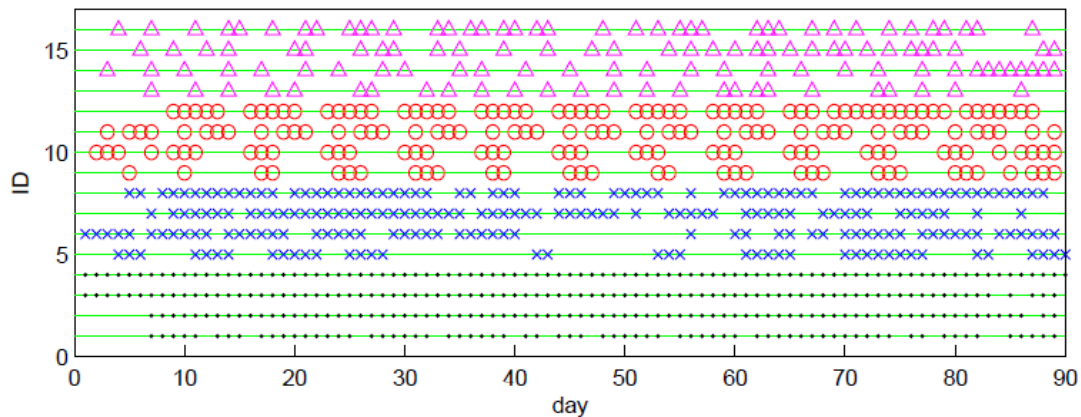




### 黄辉研究员:

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”。黄辉博士的研究方向是函数型数据分析，纵向数据分系，空间统计学，时间序列与空间点模式分析。

### 多元随机场的联合统计建模:



具有时间或空间相依结构的数据在实际应用中很常见，例如纵向数据，时空数据，图谱数据等。把数据在时间或空间上的变化轨迹看成随机曲线（随机场），就可以利用函数型数据分析技术提取数据特征。上图展现的是可卡因成瘾者的毒品使用轨迹，每个时间点上的观测数据只有 0-1 数值，但其轨迹特征包含了很多重要信息，例如周期模式和高使用量模式等。同时，每个个体可能有多个相依的非高斯轨迹变量，其交叉相依结构也包含了很多数据信息。我们近期在 JASA 上的一篇文章尝试对这种多元、非高斯的随机过程进行联合统计建模和聚类分析。其基本思想是综合利用广义线形混合效应模型，函数主成分分解和样条函数逼近等方法，来达到数据降维和模型参数估计的目的。我们的方法，将观测到的非高斯随机过程与潜藏高斯过程联系起来，把现有的函数数据分析方法扩展到任意指数族分布上。这种针对非高斯多元随机曲线的一般性联合建模方法（joint modeling）具有很强的可扩展性，也可以广泛应用在图谱和大气污染等数据分析中。另一方面，对于类似行为模式这种观测较为规则的轨迹数据，还可以从点过程的角度出发将其看成复发事件（recurrent event），这样可以降低随机效应模型参数估计过程中多重积分所带来的高计算成本。针对上图中的可卡因使用轨迹数据，我们近期在 Statistics in Medicine 上的一篇文章提出了一种利用估计方程进行联合统计建模的方法，有效的提高了运算效率并减小了模型参数估计的偏差。



**李程高级研究员：**2013 年加入北大统计中心和生命科学学院。李程博士的研究方向是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。

### 三维基因组的实验和数据分析：

李程研究组目前做实验产生 Hi-C、RNA-seq、基因组测序数据来研究染色体三维构象(3D 基因组)，并结合其他基因组数据进行整合分析和数据建模，发现癌症和干细胞领域新的生物规律或药物靶点。

在许多肿瘤和非肿瘤的人类疾病中，染色体重排 (Chromosomal rearrangement, CR) 是重要的致病原因之一。CR 事件通过破坏基因和蛋白质结构导致相关疾病的发生。此外，它们还可以通过改变染色体三维结构和基因表达而导致疾病。在这项研究中，我们使用 Hi-C 数据和 ChIP-seq 数据筛选潜在的由于三维结构改变而致病的 CR。我们设计的算法可以发现已有实验验证的和疾病相关的 CRs (如多指疾病)，这些 CR 通过破坏染色体的三维结构导致基因的异常表达。我们还发现智力缺陷可能是一个三维染色体结构改变致病的候选疾病。我们开发了一个 Web 服务器 (3Disease Browser, <http://3dgb.cbi.pku.edu.cn/disease/>)，集成和可视化疾病相关的染色体重排事件和染色体的三维结构。我们从两个染色体拷贝数变异的骨髓瘤细胞系出发，分别探索拷贝数变异 (copy number variation) 对细胞三维结构以及表达量变化的影响。发现 CNV 变化可引起染色体三维结构的差异，并且 CNV 变化的边界大多位于 TAD 的边界。CNV 也会导致基因表达量的差异，同时 A、B compartment 的变化可以引起基因表达量的变化。这些发现对理解癌症的发病机理和揭示新药物靶点有重要意义

我们在生物信息平台建设方面的进展如下：

- 1、建立和完善了 RNA-seq 和甲基化测序的分析流程。完成了 20 余个生科院和生命科学联合中心科研课题中的生物信息分析工作。
- 2、在生科院/生命中心开展生物信息和测序数据分析的技术讲座，以帮助湿实验室师生了解和使用生物信息和统计数据分析方法，以及二代测序和相关分析软件。四次讲座及暑期培训班参与的师生超过 350 人次。
- 3、利用平台论坛网站 ([www.iseq.me](http://www.iseq.me)) 普及和讨论基因组学、生物信息学、统计学和科研方法的知识和技巧，目前已发布 250 个话题。





### 林伟研究员：

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。

高维工具变量回归的正则化方法：

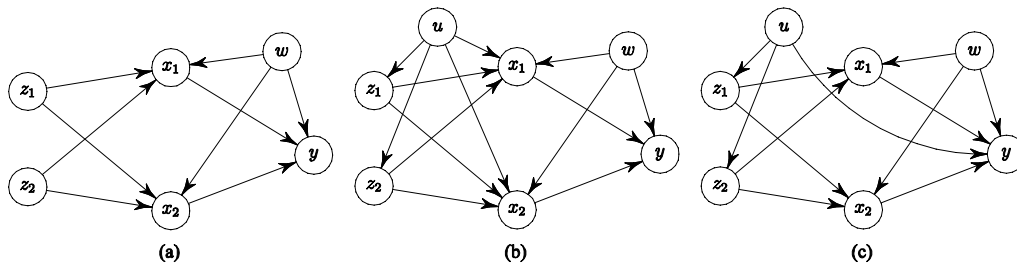


Figure 1. Causal diagrams showing the relationships between two genotypes  $z_1$  and  $z_2$ , two gene expression levels  $x_1$  and  $x_2$ , a clinical phenotype  $y$ , an unobserved phenotype  $w$  that confounds the associations between gene expression levels and the clinical phenotype, and an unobserved variable  $u$  representing possibly present population substructure. The population substructure (a) is not present, (b) affects genotypes and gene expression levels, or (c) affects genotypes and the clinical phenotype.

林伟研究员与合作者在 *Journal of the American Statistical Association* 上发表论文, 提出了高维工具变量回归的一类两阶段正则化方法, 为高维数据的因果推断提供了新的有力工具。这项研究工作的驱动问题来源于遗传基因组学。近年来, 基因表达研究的深入进展导致了众多基因标记的发现, 然而受制于混杂因素的影响, 利用这些发现破解复杂遗传性状的遗传机制的研究却仍存在重重阻力。诸多因素, 包括未能观测的变量、实验条件、环境扰动等, 都可能对基因表达施加重要的影响, 从而导致虚假的关联或对效应方向及大小的错误估计。图 1 是混杂变量与其他变量关系的一个图示。这一复杂性在高维基因组数据分析中更加突出和棘手。此外, 遗传基因组数据集的广泛获取也对基因表达和遗传变异的联合分析提出了紧迫要求, 以期对发现癌症、心血管疾病、肥胖等复杂疾病的致病基因提供关键指导。

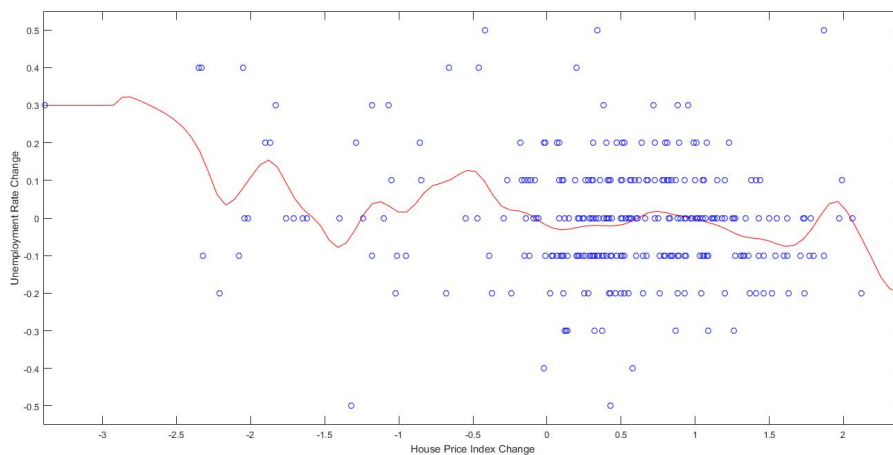
工具变量方法最早产生于计量经济学研究中并得到广泛应用, 此前在生物统计中的应用较为有限, 主要集中在流行病学研究中的孟德尔随机化问题。林伟研究组的论文指出, 工具变量模型和方法也可以用于遗传基因组数据分析, 为遗传变异、基因表达及临床性状的联合分析和混杂控制提供了一个理想框架。为了解决这一应用中协变量和工具变量的双高维度为变量选择和参数估计问题带来的方法论挑战, 该论文发展了一类两阶段正则化方法, 在选择与估计最优工具变量的同时, 可以达到识别和估计重要协变量效应的目的。该方法可以看作经典的两阶段最小二乘法的一个重要高维推广, 旨在利用稀疏性来改进两阶段的估计。该论文严格建立了这一新方法的理论基础, 并通过数值实验验证了该方法的有效性。最后, 该方法被用于一个小鼠肥胖数据集上, 找到了若干可能导致肥胖的候选基因和遗传变异。



**宋晓军研究员：**2014年在西班牙马德里卡洛斯三世大学经济系获经济学博士，同年加入北大统计中心和光华管理学院。宋晓军博士的研究方向是计量经济学理论（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。

#### 非参数模型残差的序列相关性检验：

众所周知，回归模型误差项的表现对回归函数拟合的效果是有很大影响的，如果随机误差项是独立同分布的，那么回归函数的估计和推断就可以不受误差自相关(*autocorrelation*)或者异方差(*heteroskedasticity*)的影响。然而，当随机误差项存在自相关或者异方差的情况下，回归函数的估计是一个较复杂的工作而且估计的有效性要大打折扣。如何检验误差存不存在自相关性或者异方差，或者检验误差是不是独立同分布(*independent and identically distributed, i.i.d.*)，对回归函数的估计效果以及接下来的统计推断都是至关重要的。我提出如何在非参数模型中检验误差项的独立同分布性质，即“A Nonparametric Test for Serial Independence of Errors in Nonparametric Models”。首先，传统的独立同分布检验通常是建立在参数回归模型残差(*parametric residuals*)的基础上，而非参数回归模型中关于此类的分析还很少见。其次，运用非参数模型可以避免模型误设的问题。最后，由于非参数残差(*nonparametric residuals*)的特殊性，在其基础上构建独立同分布检验所需要的分析工具也和传统参数模型不同。同时关于非参数回归模型中误差项的许多其它检验也是有待进一步研究的问题。检验非参数回归模型误差的独立同分布性质应用广泛。例如，在研究房价指数变动和失业率变动的问题上，很多人建议利用核回归方法来研究两个变量之间的动态关系，如下图所示。然而通过本文提出的检验发现，残差具有很强的相关性。这说明房价指数变动并不能很好的解释失业率变动，很容易推测出房价指数变动的滞后项也应该具有很强的解释作用。所以进一步的分析可以利用一个误差服从  $p$  阶自回归的非参数模型来解释两者的动态关系。





## 3. 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
陈大岳	副主任, 执行委员	马氏过程, 无穷粒子系统, 渗透模型, 随机游动	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
黄辉	中心全职教员	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	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	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
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	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	执行委员, 政府统计工作室召集人	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

### 三、中心要闻

#### 1. 数量金融国际会议成功举行

2015年1月13日，数量金融国际会议在北京大学光华管理学院2号楼成功举行。此次会议由北京大学数量经济与数理金融教育部重点实验室和北京大学统计科学中心联合主办，北京大学光华管理学院承办。大会组织委员会主席由光华管理学院商务统计与经济计量系年轻教员李辰旭博士担任。本次会议旨在为数量金融领域，包括数学、统计、金融、经济、计量和保险等方面的学者和专业人士提供一个交流最新研究成果的平台。近200名国内外数量金融领域的学界专家、业界人士、青年学子齐聚北大，就最前沿的理论以及实证成果进行了高水平的讨论，中国科学院院士严加安教授全程参加了此次会议。

上午的开幕式由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、北京大学统计科学中心联席主任陈松蹊教授主持，北京大学光华管理学院院长蔡洪滨教授出席开幕式并致辞，他对大会组织各方的辛勤付出以及参会者的积极参与表达了感谢，并表示组织



此类高水平的学术会议与学院的发展战略相契合。蔡院长指出随着中国经济成为世界第二大经济体，中国股票市场也跃居全球第二大股市，研究数量金融领域的相关议题对中国金融的发展是非常及时和重要的。此次会议的召开将帮助大家更好地理解金融市场现状，推

动数量金融领域的学术研究以及中国金融市场的发展。他祝愿与会者能借助大会平台展开充分的交流，取得丰硕的成果。

美国普林斯顿大学的 Yacine Aït-Sahalia 教授发表了主旨演讲。Yacine Aït-Sahalia 教授致力于研究金融计量领域的连续时间模型的估计问题，成就斐然。此次他向大家展示了运用多维自激发 jump-diffusion 模型进行投资组合最优化选择的方法。随后进入了论坛的专题报告环节。在第一个专题中，山东大学陈增敬教授就模糊情况下伯努利大数定律进行了探讨；韩国首尔大学的 Seungmoon Choi 教授与大家分享了封闭形式下的多变量非齐时性跳跃扩散过程的近似转移概率密度函数的相关研究；香港中文大学 Yue-Kuen Kwok 教授分析了离散奇异方差掉期和 timer option 定价问题。下午的报告从第二个专题开始。复旦大学马成虎教授从理论和实证两个方面讨论了 MPS-Risk-Aversion 和 Shadow-CAPM 模型；南开大学王永进教授分析了 Skew Diffusion 过程及其在金融领域的应用；中国科学院夏建明教授与大家分享了概率失真情况下的资产定价问题；微量网董事长兼量邦金融科技集团董事长冯永昌博士对量化投资做了精彩解读。在大会的第三个专题里，北京师范大学李增

沪教授探讨了 stable Cox-Ingersoll-Ross 模型估计的渐近性质；北京大学光华管理学院商务统计与经济计量系的博士生邹韬与大家分享了他用协方差矩阵回归模型动态管理金融投资组合的最新成果；北京大学光华管理学院金融学系助理教授徐江旻为大家带来了关于高频交易策略的精彩报告。

为期一天的会议很快结束了，与会者纷纷向组委会表示对本次会议高水平的学术探讨、严谨的组织工作感到满意。会议为数量金融领域的学者和专业人士提供了一个交流最新研究成果的平台，他们就最前沿的理论以及实证成果进行了高水平的探讨，思想和观点在这里碰撞和交锋。我们相信，通过各位专家学者的努力，数量金融的诸多理论和实践将会更加完善，其发展前景会更加光明！

## 2. 首届环亚太青年计量学者（YEAP）会议成功举办

2015年1月15-16日，由北京大学数量经济与数理金融教育部重点实验室和北京大学统计科学中心联合主办，北京大学光华管理学院承办的“首届环亚太青年计量学者（YEAP）会议在光华管理学院1号楼成功举行。大会组织委员会主席由受聘于统计科学中心和光华管理学院商务统计与经济计量系的年轻教员涂云东博士担任。国内外从事计量研究的著名学者与青年学者们就最新的研究成果进行了深入交流和讨论，来自全国各地的近百名青年教师和博士生参加了会议。

为期两天的会议安排了两场主旨演讲以及六组专题报告，共有20篇入选论文在会议上进行了报告和交流。光华管理学院商务统计与经济计量系的陈松蹊教授、王汉生教授、虞吉海副教授、涂云东助理教授和宋晓军助理教授分别主持了相关环节。

在15日上午的开幕式致辞中，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、北京大学统计科学中心联席主任陈松蹊教授对当今的环境问题发表了自己的独特见解，并寄希望于与会的青年计量学者，希望他们能从自己的研究领域对当下的热点问题贡献自己的力量。



Yacine Aït-Sahalia 教授主旨演讲



蔡宗武教授主旨演讲

美国普林斯顿大学的 Yacine Aït-Sahalia 教授和美国堪萨斯大学蔡宗武教授应邀分别就高频数据的主成分分析和外商直接投资问题发表了主旨演讲，引发了在场青年计量学者的热烈讨论。

15日上午，第一组专题报告以“时间序列计量”为主题，香港城市大学的韩旭老师在报告中分析了利用外部工具变量对 Structural VAR 模型的识别、估计与推断问题；中国科学



院的王霞老师向与会者介绍了时间序列的严平稳检验问题，武汉大学的谢天老师就如何进行股票回报率的预测做了报告；上海财经大学的易艳萍老师报告了协整 VAR 模型的平均估计问题。

在 15 日下午的“微观计量的识别与估计”专题中，清华大学的洪圣杰老师探讨了数据误报的识别与估计问题；上海财经大学的刘年青老师就双重拍卖的非参数识别与估计问题进行了分析；中国人民大学的徐吉良老师就有测量误差和内生性的 Single Index Model 的识别与估计问题做了报告。

第三组报告的主题是“空间和面板计量”。上海财经大学的金飞老师报告了带有 Durbin 外生变量的矩阵指数空间识别模型的估计与推断问题；中国人民大学的雷敬华老师探讨了空间自回归二元选择模型的平滑空间最大得分估计问题；同样来自中国人民大学的章永辉老师则就带有固定效应的部分线性动态面板数据模型的半参数估计结果进行了分析。



16 日上午，第四组报告以“金融计量”为主题。上海财经大学的陈强老师报告了针对高频数据的基于经验过程的遍历扩散过程的识别检验问题，武汉大学的庄额嘉老师探讨了基于四阶矩的风险度量以及其衍生的交易策略；来自西南财经大学的郭萌萌老师带来了自适应利率模型的报告；来自对外经贸大学的王芸老师分析了 Ornstein-Uhlenbeck 过程的 Mean Reversion 估计的识别问题。

16 日下午，在“假设检验”专题中，来中国自人民大学的李勇老师分享了他在贝叶斯识别检验方面的研究成果，同样来自人民大学的马骏老师带来了他在一般参数假设检验的经验似然比检验的二阶性质方面的研究，厦门大学的王学新老师展示了他在过度识别限制的一类新型检验方面的研究成果。

会议的最后一组报告探讨“微观计量的一般问题”，对外经贸大学的冯强老师分析了当参数是弱识别时如何提高置信集，华中科技大学的魏杰老师介绍了有内生性的函数系数模型中的截断分位数回归问题，复旦大学的肖志国老师与大家分享了他在带有不可忽视的缺失协变量的估计方程的半参数推断方面的最新研究。

会议最后，涂云东老师做了总结发言，他对各位嘉宾的精彩演讲，与会者的积极参与以及组委会的付出表示感谢，希冀此次大会能够为与会者今后的学术研究提供帮助。



在为期两天的会议中，参会的青年计量学者积极热烈的参与到学术讨论中，就当前的最新研究成果发布自己的见解，并互相学习借鉴。我们相信，此次会议对推动国内青年计量学者之间的学术交流以及国内的计量学科发展将起到积极的作用，并将成为一个国内以及亚太地区青年计量学者增强彼此交流，加强合作研究的最新平台。

### 3. 黄辉博士、林伟博士入选中组部第十一批“国家级重大人才青年计划”

根据中共中央组织部人才工作局及海外高层次人才引进工作专项办公室日前下发的通知，黄辉博士、林伟博士入选中组部第十一批“国家级重大人才青年计划”。

黄辉于2010年在马里兰大学巴尔的摩分校完成博士论文，2010年至2012年在耶鲁大学公共卫生学院进行博士后研究，2012年至2013年5月在迈阿密大学管理科学系进行博士后研究工作。2013年加盟北大。黄辉博士的研究方向是函数型数据分析，纵向数据分析，空间统计学，时间序列与空间点模式分析。

林伟于2011年在南加州大学获应用数学（统计方向）博士学位，2011年至2014年在宾夕法尼亚大学佩雷尔曼医学院生物统计与流行病学系进行博士后研究。2014年加盟北大。

林伟博士的研究方向是高维统计，大数据问题，因果推断，生存分析，成分数据分析，统计遗传学与基因组学。

#### 4. 第二届海峡两岸经济计量和环境统计交流与对话成功举行

2015年5月11日，由北京大学统计科学中心的陈松蹊教授和涂云东博士筹办和组织第二届海峡两岸经济计量和环境统计交流与对话在光华管理学院 K02 会议室成功举行。来自台湾中央研究院的五位学者就宏观经济预测、计量经济学和环境数据的统计学分析等主题与光华管理学院师生进行了深入的探讨交流。北京大学常务副校长刘伟教授和光华管理学院院长蔡洪滨教授分别会见了台湾中央研究院经济研究所的简锦汉所长。



简锦汉教授演讲



陈松蹊教授做总结发言

在2014年2月台湾中央研究院举行的第一届两岸宏观经济预测交流会的基础上，本次研讨会新增加备受关注的关于环境数据的建模和分析的议题。会议吸引了来自北京大学统计科学中心和光华管理学院的多位听众参加。整场会议气氛活跃，与会者就相关议题愉快交流，并于会后纷纷表示收获颇丰。

本次会议为期一天，包含三组专题共七场报告。统计科学中心的陈松蹊、虞吉海、涂云东和宋晓军老师参加并主持了专题讨论。

会议于上午九点整准时举行。首先，北京大学光华管理学院副院长金李教授致辞。金李教授对来自台湾的学者表示热烈欢迎并表达了对会议议题的强烈兴趣和关注，希望本次交流能够促进两岸在经济和环境分析方面的合作。之后台湾中央研究院经济研究所所长简锦汉致辞，他回忆了去年在台湾成功举办的第一次交流会，同时希望本次会议能够取得更大的成功。北京大学荣誉教授刁锦寰先生也发言鼓励与会学者广泛交流，并相信本次交流会一定能促进双方学术研究的共同进步。

上午的第一个专题讨论包含两场报告。首先，简锦汉教授讨论了教育对人体健康的影响。接下来，台湾中央研究院经济研究所的周雨田教授讨论了如何利用带异常点的近似因子模型做经济预测。上午的第二个专题讨论包含两个环境统计的报告。台湾中央研究院统计科学研究所的黄景祥教授探讨了空气中颗粒物污染与心血管疾病的关系。之后光华管理学院的陈松蹊教授讲述了用统计方法对北京城区的PM2.5等环境数据的研究。

下午的专题讨论包含三场报告。第一场报告由台湾中央研究院经济研究所的陈宜廷教授给出，他讲述了在分组数据下估计和检验收入分布的方法。之后光华管理学院商务统计与经济计量系的涂云东博士提出了自适应货币数量理论并展示其在通货膨胀预测中的优秀



表现。最后，台湾中央研究院经济研究所的蔡文祯教授报告了线性混合频率回归模型的估计和应用。

最后，陈松蹊教授做了总结发言。他对各位嘉宾的精彩演讲、与会者的积极参与以及组委会的辛勤付出表示感谢，并希望两岸对话持续进行，推动两岸交流平台的不断完善，促进两岸学术界在经济学、统计学乃至更广泛范围内的合作与交流。

## 5. 第八届中国 R 语言会议暨 2015 北大光华数据与价值论坛成功举办



6月6日-7日，第八届中国 R 语言会议暨 2015 北大光华数据与价值论坛在北京大学成功举办。本次会议由北京大学光华管理学院和统计之都联合主办，北京大学商务智能研究中心、北大光华—奥迪管理研究中心、北京大学统计科学中心协办。

大会开幕式由会议主席、北京大学光华管理学院商务统计与经济计量系在读博士生朱雪宁主持，北京大学光华管理学院院长蔡洪滨教授代表主办方发表了致辞，蔡院长从商学院的视角对数据科学的重要意义做了独特的阐述，并对辛勤工作的会议志愿者表达了感谢。



随后，本次会议的主要推动者和组织者、北京大学光华管理学院商务统计与经济计量系系主任王汉生教授也发表了致辞。王汉生教授不仅学术成果斐然，更有浓浓的数据科学情怀，为推广 R 语言以及数据科学知识做出了不可磨灭的贡献。



在随后的主题演讲中，美国艺术与科学院院士、加州大学伯克利分校统计系系主任郁彬教授出场，她展示了一个当下正在进行的项目，为大家揭示了基因数据背后的规律。由于基因数据的复杂性和高维性，对其进行有效的数据分析以发现其中的价值显得格外困难，往往需要数据科学家和生物学家的通力协作。百度大数据实验室负责人、美国新泽西大学统计系教授张潼博士在报告中介绍了百度处理大数据的一些案例，技术积累以及挑战。白硕先生在演讲中

对未来社会中知识服务的前景做了展望，并提到通过深层自然语言处理技术实现的知识引擎，将是知识服务和智能终端共同的核心组件，值得业界高度关注。袁晓如教授认为，拥有大量的数据并不等于获得相应数据的价值，复杂数据对发展相应的分析工具提出了新的挑战和要求。在报告中，冯永昌先生提出用股指期货的日内程序化交易策略和做空波动率的期权组合形成风险对冲策略，即 Risk-Neutral 策略，模拟业绩显示复合策略大幅提升夏普比

率，平滑曲线走势。这是用股指期货管控量化交易风险的一个实验性案例。苏萌博士现任百分点董事长兼 CEO。他在报告中阐述了大数据产业链的全球生态格局和大数据的技术与应用，同时展示了大数据的商业建模应用实例。叶信岳教授来自肯特州立大学，在报告中，叶教授对空间不平等数据的分布及建模做了详细阐述，对空间数据的实际应用提出了自己的看法。王汉生教授在报告中提出基于社交网络的网络结构数据可为实现快速准确征信提供支持，并通过实际案例生动详细的做出了分析和探讨。

在为期两天的会议里，参会者积极参与到数据科学的讨论中，就当前的最新研究成果发表自己的见解，并互相学习借鉴。今年是中国 R 语言会议举办的第八个年头，参会人数突破了 3000 人，创下历届之最，使本次会议成为亚太地区规模最大的数据科学盛会之一。我们相信，本次会议将对 R 语言以及数据科学知识的推广和发展起到积极的作用，也将成为国内以及亚太地区数据科学领域的学者和行业人士增强彼此交流，加强合作研究的重要平台。

## 6. “大数据时代的高维统计”国际青年学者会议暨夏季短期课程

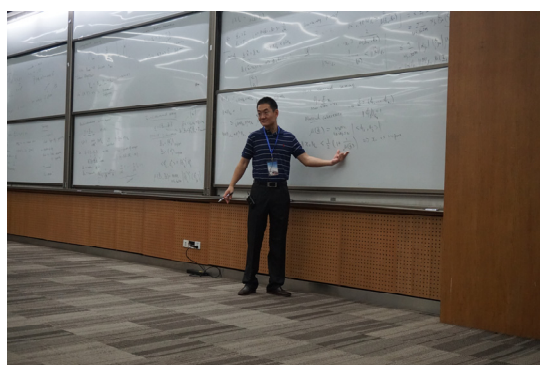
“大数据时代的高维统计”国际青年学者会议暨夏季短期课程于 2015 年 6 月 25 - 28 日在北京大学举行。此次会议暨短期课程由北京大学统计科学中心和北京大学数学科学学院概率统计系联合主办。



蔡天文教授致开幕词



范剑青教授做主旨报告



袁明教授短课



周慧斌教授短课



高维统计是当前统计学研究最活跃的前沿领域之一，与统计机器学习、压缩感知、大数据分析等数据科学分支有着深刻联系和高度交叉。为应对大数据时代为统计理论、方法及应用带来的重大挑战与机遇，此次会议以“大数据时代的高维统计”为主题，旨在凝聚海内外从事高维统计相关领域研究的新生力量，为近年来在这一领域崭露头角的青年学者提供一个交流合作与展示最新研究成果的平台，推动高维统计理论与实践的蓬勃发展，促进与互联网经济、精准医疗等大数据时代重要应用问题的紧密结合。

此次会议邀请到美国普林斯顿大学范剑青教授做主旨报告，以及海内外著名高校 20 余名青年学者（获得博士学位均不超过 7 年）做邀请报告。会议议题十分丰富，涵盖高维统计（从广泛意义下解释）的各个研究方向，包括变量筛选与模型选择、大协方差矩阵和图模型估计及检验、高维时间序列与数据流、高维聚类与判别分析、因子模型与主成分分析、大数据下的抽样与众包等等。我们还邀请到美国威斯康星大学麦迪逊分校袁明教授和美国耶鲁大学周慧斌教授，在此次会议前开设分别题为“稀疏学习”和“图模型与网络分析中的最优估计”的夏季短期课程，针对青年教师、博士后及研究生，讲授高维统计领域的若干前沿问题和最新进展，本次活动吸引了来自全国的 150 多人前来参加。

## 7. 概率论暑期系列学术活动



为了促进北京大学概率论学科的发展，扩大北京大学概率论学科在国内外的影响力，加强交流，活跃学术气氛，发掘人才，开阔视野，2015 年暑期，北大概率论教研室举办了系列的学术活动。这一系列的活动包括 7 月 6 日至 7 月 8 日举办的“2015 年北京大学青年概率论坛”和 7 月 9 日至 7 月 24 日的两门短期课程。

“2015 年北京大学青年概率论坛”邀请了来自美国、加拿大、英国、瑞士、新加坡等国外知名大学以及国内包括香港、澳门在内的共计 20 余位优秀和活跃的青年概率学者做邀



请报告。这些报告涉及高度奇异随机偏微分方程的正则性理论、高斯自由场、交互作用的粒子系统，随机树以及概率论在生物、物理、化学、流体力学和金融保险中的应用等前沿领域，充分反映了当前国际概率论发展热点和趋势。这些报告也引发了与会者热烈的讨论。与会者普遍认为这次青年论坛非常成功，富于朝气与活力，为年轻学者提供了一个相互认识相互了解，交流切磋的平台，也使得年轻学者能够深入了解北大。

紧接着论坛，为研究生开设了两门概率论前沿研究的短期课程，分别邀请了美国明尼苏达大学的胡文清博士讲授“Stochastic Fluid Mechanics”和英国 Bath 大学的 Kyprianou 教授讲授“Levy Processes and Self-Similar Markov Processes”。这些课程的听众不但包括北大、中科院、北师大等单位概率论专业研究生和教师，而且还吸引了基础数学，应用数学等其它方向的老师和研究生前来听讲。在课程间隙，还组织了任课老师和研究生的讨论班，就课程的习题，难点和进一步发展进行了深入的探讨，取得了良好的效果。

这次暑期概率论的系列学术活动受到北京大学统计科学中心，国家自然科学基金，科技部 973 计划，高等学校数学研究与高等人才培养中心的资金资助，得到了北大统计中心，北大数学学院在人员和场地等多方面的支持。暑期学术活动圆满成功。

## 8. 王汉生教授获得国家杰出青年科学基金项目资助

在 2015 年国家杰出青年科学基金项目中，统计科学中心执行委员会委员，光华管理学院商务统计与经济计量系联席系主任王汉生教授申报的“高维复杂数据的理论与应用”榜上有名。

王汉生教授多年来致力于高维数据分析，统计学在电子商务领域的应用等研究，尤其关注网络数据和位置轨迹数据的统计分析。他 1998 年北京大学数学学院概率统计系本科毕业，2001 年美国威斯康星大学麦迪逊分校统计系博士毕业。现任北京大学光华管理学院商务统计与经济计量系教授，博士生导师，系主任；北京大学商务智能研究中心主任；博雅立方科技有限公司首席科学家；微信公众号“狗熊会”创始人。近年来，他发表英文学术论文共计五十余篇，中文论文近二十篇。合著英文专著 1 本，独立完成中文教材 2 本，先后担任多个学术刊物副主编(Associate Editor)，2014 年当选美国统计协会会员，此次获得国家杰出青年科学基金项目资助正是对他在统计学界孜孜以求多年并取得丰硕研究成果的极大肯定。

## 9. 统计科学中心成功举办 2015 年暑期夏令营



2015 年 7 月 10 日至 12 日，北京大学统计科学中心举办了第二届优秀大学生夏令营活动，来自中国人民大学、南开大学、西安交通大学、武汉大学、厦门大学等十几所高校的 25 名三年级本科生参加了本次夏令营。

此次夏令营的活动包括科研文献阅读理解、笔试与面试考核、统计中心老师介绍各课题组的科研情况、师生聚餐、夏令营学员的英文自我介绍等环节。

黄辉老师在开幕式上给参加夏令营的同学介绍了统计科学中心的基本情况。统计中心联席主任陈松蹊老师做了关于 PM2.5 大气污染统计建模的特邀科研报告，引起了

学员们的浓厚兴趣。美国华盛顿大学生物统计系的周晓华教授介绍了他的科研兴趣与成果。数学科学学院金融数学系系主任吴岚老师介绍了她的研究领域和成果。统计中心的其他全职教员也介绍了自己的科研方向，并与同学们进行了深入的对话和交流，让学员们对统计中心和国际前沿的统计学研究方向有了更加全面的了解。在这次夏令营中，统计中心的老师们选拔了表现突出的几名学员，并推荐保送北京大学直博研究生项目。



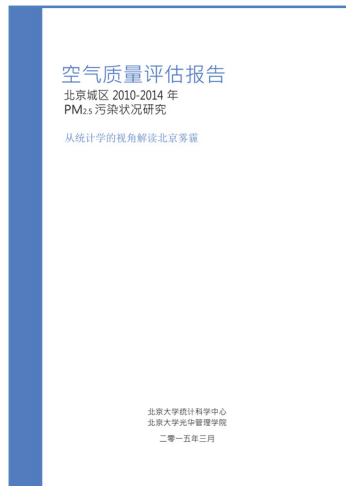
## 10. 《空气质量评估报告：北京城区 2010-2014 年 PM<sub>2.5</sub> 污染状况研究》发布

2015 年 3 月 24 日，由陈松蹊教授率领的北京大学统计科学中心和光华管理学院环境统计课题组正式发布了《空气质量评估报告：北京城区 2010-2014 年 PM<sub>2.5</sub> 污染状况研究》。报告旨在为如何治理中国大气污染问题提供一个数据的视角和统计学上的分析判断，一经发布便得到了各大媒体的深入报道，引起了社会公众的广泛关注。

大气污染问题一直是关乎经济和社会科学发展的关键问题，近年来更是在全国范围内得到了越来越多的关注和讨论，尤其是细颗粒物（PM<sub>2.5</sub>），随着其污染问题的凸显和加重，如何治理 PM<sub>2.5</sub> 污染逐渐成为制定经济和社会发展政策中的热点问题。2013 年 1 月 1 日，环保部正式将 PM<sub>2.5</sub> 列入空气监测指标中。2013 年 1 月，北京频繁出现极端雾霾天，PM<sub>2.5</sub> 污染受到广泛关注。2013 年 9 月 10 日，国务院印发《大气污染防治行动计划》，其中提到，到 2017 年，京津冀地区细颗粒物浓度比 2012 年下降 25% 左右。在此背景下，北京大学统计科学中心和光华管理学院的八位老师和同学开始收集和分析北京地区的 PM<sub>2.5</sub> 数据。

这份《空气质量报告》概括了团队一年来的研究成果，主要涉及了六个方面，包括：美国驻华大使馆与环保部数据的一致性验证、PM<sub>2.5</sub> 污染的基本统计分析、风向和风力对 PM

2.5 污染的影响、2010 年至 2014 年 PM2.5 年均值浓度及分位数浓度的公平比较、APEC 效应、冬季供暖效应，分别得到了如下结果与结论。



在数据方面，报告采用了美国驻华大使馆 2010 年到 2014 年共五年的 PM2.5 逐小时浓度数据和北京首都国际机场（ZBAA）自 2010 年至 2014 年的逐小时气象数据，考虑的气象变量有气压、温度、相对湿度、露点温度、风向、风速和降水。为了相互验证美国大使馆与环保部数据的可靠性，报告比较了从 2014 年 5 月 1 日到 2014 年 12 月 31 日美国驻华大使馆与环保部下属六个在京站点的逐小时 PM2.5 观测值，发现二者的数据高度一致。

在污染的基本统计方面，报告首先根据 PM2.5 小时浓度将空气质量分为三级：优良状态（PM2.5 浓度不超过 35 微克 / 立方米）、污染状态（PM2.5 浓度大于 35 微克 / 立方米）、严重污染（PM2.5 浓度大于 150 微克 / 立方米），并在此基础上，计算了北京过去五年 PM2.5 的三种空气质量状态的持续时间。结果显示，北京地区每年约有 23% 的优良天气，55% 的一般污染天气，22% 的严重污染天气。这三个百分比在过去五年中的变化不大。

在风对 PM2.5 浓度的影响方面，报告指出：（1）北风对北京城区的污染有很强的清洗作用，而南风会把北京以南的污染物传输过来，加剧北京城的空气污染。（2）如果没有一级以上北风，北京的 PM2.5 浓度在优良水平状态的平均持续时间将从 20 小时大幅下降到 4.7 小时。这说明北京自身排放的污染是不可忽视的。

在对过去五年 PM2.5 年均值浓度及分位数浓度的比较方面，报告通过现代统计学方法，首次给出了一个公平地比较不同年份的 PM2.5 浓度的方法。这一方法扣除气象因素的影响，使不同年的 PM2.5 可以公平比较。报告通过分析去除气象因素影响后的均值及分位数指标发现，与 2012 年相比，2013 年和 2014 年北京城区的 PM2.5 污染水平没有显著改善。

在以上结论的基础上，报告从政策分析的角度，采用了准实验研究的统计学方法，分析了 APEC 期间大力度减排措施对污染减轻的作用。报告通过分析具体量化的统计学指标发现，“APEC 蓝”是北京及周边地区超大力度减排的结果，但减排后平均 PM2.5 仍有 51.5 微克/立方米。这说明在现有的能源消耗模式下，北京地区获得持续性达标空气仍任重道远。

进一步，报告分析了冬季供暖对 PM2.5 的影响，给出了供暖期与非供暖期的 PM2.5 调整浓度，由并此得出，只通过减少冬季供暖的污染不能实现北京的 PM2.5 浓度比 2012 年下降 25% 的目标。

最后，报告综合以上分析结果指出，在现有的减排力度与减排模式下，北京很难达到《大气十条》所订下的 2017 年的两个减排目标（京津冀地区细颗粒物浓度比 2012 年下降 25% 左右，北京市 PM2.5 的年均浓度控制在 60 微克/立方米）。要实现这两个目标，京、津、冀、鲁、豫必须进行能源消费结构和产业结构的变革性调整，及环境法规的严格执行。



## 11. 北京城区 PM2.5 污染状况的研究在《皇家学会会刊》发表

日前,由陈松蹊教授带领的光华管理学院和统计科学中心的环境大数据研究团队关于北京城区 PM2.5 污染状况的研究 (Assessing Beijing's PM2.5 Pollution: Severity, Weather Impact, APEC and Winter Heating) 被英国皇家学会会刊《Proceedings of the Royal Society A》正式发表。

文章的第一作者是光华商务统计系博士生梁萱,其它作者是邹韬(光华博士生)、郭斌(毕业于光华,川大助理教授)、李硕(光华博士生)、张豪哲(Iowa State University 统计系博士生)、张澍一(光华博士生)、黄辉(数学学院概率统计系,统计科学中心助理教授)、陈松蹊(光华,统计中心)。陈松蹊和黄辉为共同通讯作者。

背景:我们观测到的 PM2.5 实时浓度数据受气象因素的影响很大。这就导致一个地方的污染排放实际上增加了,但由于有利的气象扩散条件,观测的 PM2.5 反而可能会更低。同理,由于在不同年份,不同城市气象条件不同,导致 PM2.5 浓度变化有差异,直接比较原始 PM2.5 的浓度会不科学甚至不公平。中国的大气污染防治需要一个科学的空气质量评估方法。该文章发现北京 PM2.5 月度数据近 80%的变化(variation)是由于气象因素的变化所导致。

环境大数据研究团队的文章利用统计学方法分析过去五年北京 PM2.5 浓度与气象变量的关系。提出了一个公平地评估不同年份 PM2.5 浓度是否有显著改善的方法。该文章提出应该去除掉气象变量的影响,对原始浓度进行调整以反映真实的城市污染、排放情况,从而为政府环保部门提出合理的、科学的指导方案。文章中提出的统计学方法可以广泛地用于其它城市和地区的空气质量评估。

该文章通过学习、分析过去五年的污染和气象数据,发现在去除气象因素的影响后,2013年和2014年北京城区的 PM2.5 污染水平与2012年相比并没有明显改善。且2013年和2014年的 PM2.5 年平均浓度比2012年分别增加了10.6%和5.9%。

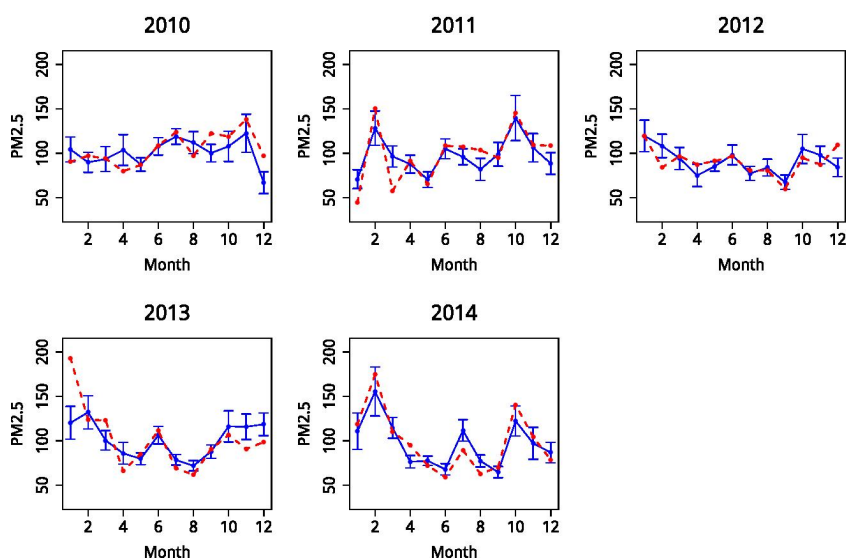


Fig.1: 北京 2010-2014 年 PM2.5 月均值原始浓度(红色)及调整后月均值浓度(蓝色及其相应的 95%置信区间)。

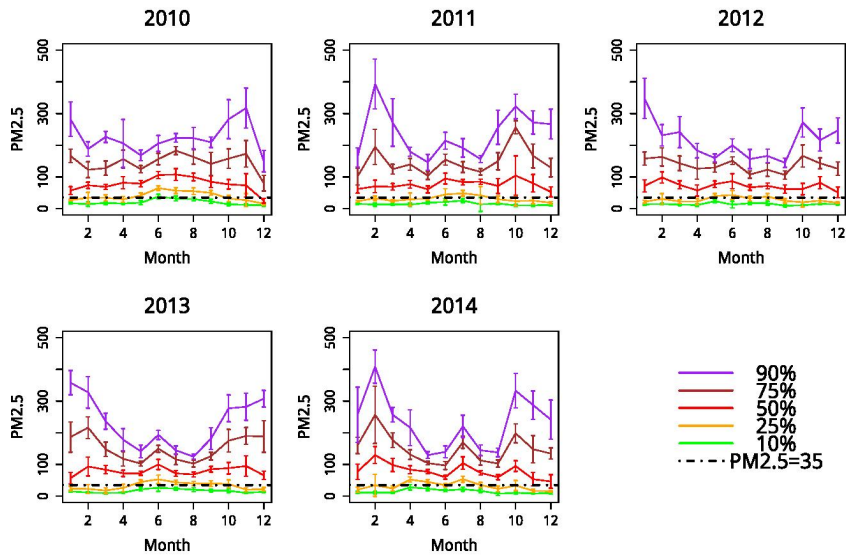
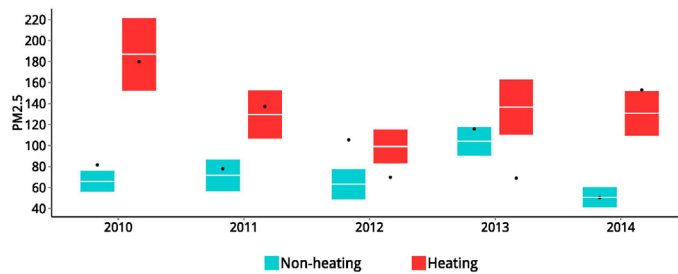


Fig.2: 北京 2010-2014 年 PM2.5 调整后浓度的月分位数。

利用上述调整方法，该文章还分析了 2014 年 APEC 期间政府减排措施的控污效应及 2010-2014 年北京冬季供暖效应。结果发现，“APEC 蓝”确实是北京及周边地区采取了极大减排措施的结果。冬季供暖期比非供暖期的 PM2.5 浓度平均要高出 50%多。如果不使用文章中的方法进行调整，供暖效应在某些年份很难体现。

(a) 十一月冬季供暖效应



(b) 三月冬季供暖效应

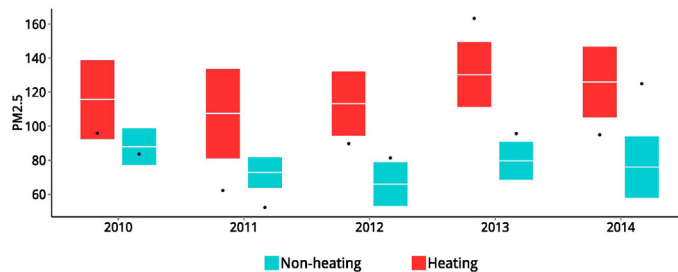


Fig.3: 冬季供暖效应: (a) PM2.5 在十一月非供暖期 (蓝色) 和供暖期 (红色) 的原始均值浓度 (黑色实点)、调整后均值浓度 (白色实线) 及其 95%置信区间。(b) PM2.5 在三月份供暖期 (红色) 和非供暖期 (蓝色) 的原始均值浓度 (黑色实点)、调整后均值浓度 (白色实线) 及其 95%置信区间。

文章最后提到,北京若要达到《大气十条》所订下的PM2.5浓度标准,即2017年相比2012年减少25%和年平均浓度不高于60微克/立方米的目標,政府需要实施更有效的减排措施并且真正执行环保法规,否则PM2.5的污染水平很难大幅度降低。

该研究得到了北京大学统计科学中心的资助。



## 12. “大数据驱动的管理决策模型与算法”重点项目获批

陈松蹊教授带领的主要由统计中心教员组成的研究团队成功申请了国家自然科学基金重点项目“大数据驱动的管理决策模型与算法”,研究期限为:2016年01月--2020年12月。该项目团队主要参与者有南开大学的王永进教授,北京大学的李程研究员、虞吉海教授、席瑞斌研究员、林伟研究员、涂云东研究员和宋晓军研究员,西南财经大学的讲师兰伟。

该项目旨在利用创新的统计学方法和可行的算法对在金融、经济、医疗等领域中出现的基于互联网、电子商务等新的数据类型进行分析。该项目包含了四个目前国际大数据研究的前沿问题,分别为:(1)大数据统计推断理论,(2)多类型数据的整合分析算法,(3)金融“信用”与“利率”的实时计量分析和(4)医疗健康大数据挖掘的基本方法。这四方面的研究更新了传统的管理学研究方法,其所提出的大数据的统计推断理论和多类型数据的整合分析算法可以直接转化为能应用到大数据金融、医疗健康等领域的具体算法和数据挖掘的模式,提升这些领域科学利用大数据的实践能力。同时该项目所提出的大数据分析模型和算法可以应用到工商管理 and 经济管理方面的量化决策中,促进大数据驱动下管理科学决策理论和方法的发展。

## 13. 中心教员社会服务情况,包括对各级政府部门提供顾问、咨询等服务情况

王汉生教授: 奥迪+奇虎 360, 北大光华新年论坛, 奥迪专场; 奥迪, 横向课题, 高端车客户价值分析; 腾讯, workshop, 大数据与隐私合规; 量帮, 横向课题, 量化投资。

席瑞斌研究员: 高强, 复旦大学中山医院, 肝癌基因组的异质性问题; 席建忠, 北京大学工学院, 单细胞测序技术。



张俊妮副教授: 中国信达资产管理公司, 金融不良资产估值定价模型研究; 中国青年报和光华管理学院合作的”青年公众 CSR 认知调查” 研究项目; 腾讯, workshop, 大数据与隐私合规 (参与)。

林伟研究员: 宾夕法尼亚大学佩雷尔曼医学院, 微生物组与宏基因组。

李程研究员: 孙育杰 (北京大学生科院, 染色质三维结构成像。结合染色体构象捕捉实验的数据分析和高精度荧光成像技术, 确定细胞核内转录工厂的构成。); Peter Sicinski (哈佛大学医学院, 筛选小 RNA 作为药物靶点。通过实验筛选抑制细胞周期蛋白表达量的小 RNA, 并用基因组表达谱预测癌症细胞对小 RNA 作为药物的敏感性。); 季红斌 (中科院上海生科院, 肺癌靶向药物的抗药性机制。通过实验产生对靶向药物抗药的肺癌细胞系, 发现可能的抗药基因和机制)。

虞吉海: Spatial Economic Analysis 的联合主编; Regional Science and Urban Economics 的副主编。

杨静平: 《中国国债管理战略计量分析课题巴西随机模拟模型研究》, 中央国债登记结算有限责任公司, 项目负责人; 《信贷资产证券化组合信用风险模型理论研究》。中债资信评估有限责任公司, 项目负责人。

黄辉研究员: 官永涛, 迈阿密大学商学院, 空间流行病的数据合并技术; 沈晔, 佐治亚大学公共卫生学院, 纵向数据与复发事件的联合统计建模; 王学钦, 中山大学华南统计中心, 李扬, 中国人民大学统计学院, 脑电图的统计建模与推断; 郭巍, 中国疾病预防控制中心, 艾滋病的时空风险分析; 王自发, 中科院大气物理所, 大气污染的时空分析

艾明要教授: 主持完成总装备部陆军武器装备体系国防科技重点实验室研究课题“面向武器装备体系仿真试验的数据分析关键技术研究”; 完成北京理工大学委托国防预研合作课题“CT 扫描含能材料缺陷的显著影响因素筛选方案”; 参与国家体育总局信息中心“体育产业调查统计方案设计”咨询; 三个国际期刊 JSPI、SPL、STAT 的 Associate Editor。

#### 14. 中心教员博士毕业生走向

姓名	专业	工作单位	指导教师
郭斌	统计学	四川大学经济学院	陈松蹊
黄丹阳	统计学	人大统计学院	王汉生
马莹莹	统计学	北航管理学院统计系	王汉生
高扬	商务统计	北京工业大学经济管理学院师资博士后	王明进
贾晨	概率论与数理统计	北京计算科学研究中心	蒋达权
张蕊	概率论与数理统计	首都师范大学	任艳霞
李康	概率论与数理统计	中国人民银行营业管理部	艾明要
罗鹏	概率论与数理统计	深圳大学	耿直
郑嘉寅	概率论与数理统计	九州证券有限公司深圳分公司	何书元
房厦	统计学	首都经济贸易大学	杨静平
李禄俊	应用数学	中金基金管理有限公司	杨静平
王艺舒	数学	青岛大学	邓明华

## 四、交流访问

### 1. 北京大学名誉教授刁锦寰教授北大课程

国际著名统计学家北京大学名誉教授刁锦寰教授于4月27日至6月15日访问北京大学。访问期间于5月8日开始在光华管理学院1号楼开设以“Outlier Detection, Intervention Analysis, Vector ARIMA model, 多元时间序列的实例应用, Seasonal Adjustment”为主题的统计学课程。



国际著名统计学家刁锦寰教授长期从事研究与教学工作，曾先后在威斯康辛大学、哈佛大学和芝加哥大学任教，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，北京大学名誉教授。

刁锦寰教授在学术与研究领域中享有盛誉并多次获奖，曾于1961-1962年获Ford基金论文奖，1993年获台湾主计处卓著奖，2001年获华盛顿统计学会和商业协会授予的Julius Shiskin奖，2001年荣获美国统计学会授予的Samuel S. Wilks奖章。刁锦寰教授在1973年就成为美国统计协会荣誉会员；1974年为数学统计协会荣誉会员；1976年为台湾中央研究院院士；1980年为国际统计协会会员。

### 2. “贝叶斯推断”短期课程顺利完成

2015年5月5日至6日，由北京大学统计科学中心与北京大学光华管理学院商务统计与经济计量系联合举办的“贝叶斯推断”短期课程在光华管理学院圆满完成。

本次课程由国际著名统计学家Alan E Gelfand教授主讲。Alan E Gelfand教授是杜克大学James B Duke统计科学教授，发表过270余篇论文，因贝叶斯计算、贝叶斯推断、空间数据统计建模等方面的研究建树而在全球范围内拥有广泛且深远的影响力。Alan E Gelfand教授是美国统计学会（American Statistical Association）、国际数理统计学会（Institute of Mathematical Statistics）以及国际统计学会（International Statistical Institute）会士，曾任国际贝叶斯分析学会（International Society for Bayesian Analysis）主席。

本次短期课程面向国内各高校研究生、青年教师免费开放，授课现场，近200名来自全国各地的学者、学生聆听了Alan E Gelfand教授的课程，原本额定80人的教室气氛热烈，蔚为壮观，以至于第二天不得不更换至更开阔的教室。

本次短期课程的内容既包含贝叶斯推断的基础内容，也包含了贝叶斯统计的前沿研究问题，与会者在课上和课下与Alan E Gelfand教授展开了热烈而深入的互动。短短两天的课程，Alan E Gelfand教授讲授了包括贝叶斯定理、先验分布与后验分布、模型比较与模型选择、贝叶斯推断与传统推断的差异等贝叶斯推断的基础知识，以及包括随机效应模型、缺失数据模型、潜在变量模型、动态模型等特例的分层模型；并讲授了多种贝叶斯计算方法，包括MCMC算法、粒子滤波算法、INLA、ABC、路径抽样（Path Sampling）等；还结合多种数

据实例讲述了对空间数据的贝叶斯建模方法。内容之丰富，可谓盛宴。除了渊博的学养，Alan E Gelfand 教授激情且不失风趣的授课风格也给与会者留下深刻的印象。



### 3. 蔡天文教授来访统计科学中心并做报告

国际著名统计学家、2008年COPSS奖得主，宾州大学沃顿商学院 Dorothy Silberberg 讲席教授、北京大学短期“国家级重大人才计划”专家蔡天文 (Tony Cai) 教授近日访问统计科学中心，并于 5月26日下午在理科一号楼 1114 教室作了题为“Robust Detection and Recovery of Structured Signals”的精彩学术演讲。



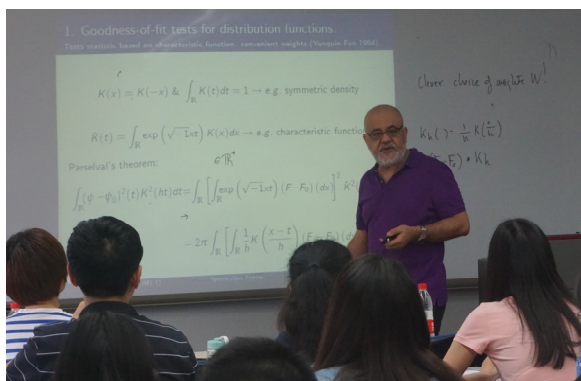
在演讲中，蔡教授分享了他在信号还原与噪声处理方面的最新研究成果和统计学思想，提出了利用 Median Coupling 的概念发展出局部中值变换 (Local Median Transformation) 技术，为非高斯信号处理这一领域提供了一种新的思路和稳健方法。传统统计学中，许多方法和理论都依赖于高斯假设，这是因为高斯分布是一种被充分研究的概率分布，并且是中心极限定理的核心内容。然而现实当中我们观察到的数据或者

信号，往往存在非高斯的测量误差或者噪声，这使得信号还原十分困难。蔡教授提出的局部中值变换技术，可以将观测信号转换成标准的高斯分布，并利用已有的针对高斯信号的技术对其进行还原。这一方法不仅可以稳健的找到真实的信号，同时又具有很好的大样本性质。蔡教授在演讲中不仅介绍了该方法的理论性质，还介绍了其在拷贝数变异方面的成功应用。



蔡天文教授的演讲吸引了北大数学学院、光华管理学院以及其他众多北京高校的师生前来聆听。演讲结束后蔡教授还与听众积极互动，进行了热烈的讨论。

#### 4. “计量经济学中的模型设定检验”短期课程



“计量经济学中的模型设定检验”短期课程由来自马德里卡洛斯三世大学 Miguel A. Delgado 教授主讲。该课程吸引了北京地区多所高校的青年计量经济学者。Miguel A. Delgado 教授强调该课程是对计量经济学中重要假设检验问题的一次系统梳理工作，并且成功的利用一个统一的框架来解释和联结假设检验中的各种检验方法。

#### 5. 徐歆怡博士访问统计科学中心五个月

美国俄亥俄州立大学副教授徐歆怡博士于今年2月至7月来统计科学中心访问五个月，在概率统计系开设了《贝叶斯统计》课程，热情介绍了数理统计学在各个领域的应用，深受同学们欢迎。访问期间她还于4月2日在理科一号楼1114教室做了题为“Shape-constrained Semiparametric Additive Stochastic Volatility Models”学术报告。

徐歆怡2001年毕业于中国科学技术大学，2005年在宾夕法尼亚大学获得博士学位，她的研究领域主要包括贝叶斯分级模型，统计决策理论，高维数据分析，生物统计学、金融和经济方面的统计应用。

#### 6. 丹麦哥本哈根大学 Theis Lange 博士作演讲

哥本哈根大学生物统计部副教授 Theis Lange 博士近日访问统计科学中心，并于12月10日下午在理科一号楼1114教室作了题为“Recent developments within causal inference based effect separation”的精彩学术演讲。Theis Lange 博士2008年毕业于哥本哈根大学，此后从事生物统计、流行病学、计量经济学等领域的研究工作，有许多优秀的研究成果。



在演讲中，Theis Lange 博士分享了他在因果推断方面的最新成果和统计思想，提出了利用不同的因果路径来做 Mediation Analysis 的概念。这一研究的目的在于在流行病和社会科学实践当中，更好的理解事物发展的机制，从而合理的提出干预措施。Theis Lange 博士提出了一系列的自然效应模型，将 Mediation Analysis 放入传统的回归模型当中，并给出了模型参数的解释方法。此外，Theis Lange 博士还介绍了相关研究的 R 程序包的使用。

此外，Theis Lange 博士还介绍了相关研究的 R 程序包的使用。

Theis Lange 博士的演讲吸引了北大数学学院、生命科学院以及其他众多北京高校的师生前来聆听。演讲结束后 Theis Lange 博士还与统计中心的老师、同学进行了热烈的讨论。

## 7. 统计中心 2015 年来访人员名单

2015 年统计科学中心共接待海内外访问学者 53 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
Tony Cai, University of Pennsylvania	2015-01-14/17
Yin Xia, UNC-Chapel Hill	2015-01-15
Shen Weining, The University of Texas MD Anderson Cancer Center	2015-02-24
Xinyi Xu, Ohio State University	2015-02/07
丁剑博士, 芝加哥大学	2015-03-09
Juxin Liu, Department of Mathematics and Statistics, University of Saskatchewan	2015-03-13/27
Johannes Ledolter, University of Iowa	2015-03-23
Peter M. Robinson, London School of Economics	2015-03-25
Jinyuan Chang, The University of Melbourne	2015-03-31
Junhui Qian, Shanghai Jiao Tong University	2015-04-08/10
Sung Yong Park, Chung-Ang University	2015-04-15/17
葛颢, 北京大学	2015-04-20
Myung Hwan Seo, Seoul National University	2015-04-22/24
邹长亮, 南开大学	2015-04-30
Prof. Fang Yao, University of Toronto	2015-05-10/12
Mengmeng Guo, Southwestern University of Finance and Economics	2015-05-20/22
宋永生副研究员, 中科院数学院	2015-05-25
Tony Cai, University of Pennsylvania	2015-05-26
Zhongjian Lin, Emory University	2015-05-27
胡涛, 首都师范大学数学科学学院	2015-05-28
Prof. Piotr Garbaczewski, University of Opole, Poland	2015-06-01
Kai Zhang, UNC-Chapel Hill	2015-06-03
George C. Tiao, University of Chicago	2015-04-27/06-15

Yili Hong, Virginia Tech	2015-06-03/07
Bo Lu, Ohio State University	2015-06-11
Jong-Myun Moon, University of College London	2015-06-16
Yi Huang, University of Maryland Baltimore County	2015-06-19
Zheng Xu, University of North Carolina at Chapel Hill	2015-06-24/28
王永雄, 斯坦福大学	2015-06-26
Xiao Wang, Purdue University	2015-06-28/30
范剑青, 普林斯顿大学	2015-07
肖益民教授, Michigan State University	2015-07-13
Jingchen Liu, Associate Professor, Columbia University	2015-07-16
Simon Tavaré, Cancer Research UK Cambridge Institute & DAMTP, University of Cambridge	2015-08-10/13
Nancy Reid, Department of Statistical Science, University of Toronto	2015-08-12
Chi-Lun Cheng, 台湾中央研究院	2015-09-17
Peisong Han, University of Waterloo	2015-09-24
Wei Lin, Capital University of Economics and Business	2015-09-24
罗德军副研究员, 中国科学院数学院	2015-09-28
Ke Yuan, University of Cambridge	2015-10-12
郑泽敏, 中国科学技术大学	2015-10-13
Lijian Yang, Soochow University	2015-10-14/15
Jun Yu, Singapore Management University	2015-10-15
唐年胜教授, 云南大学数学与统计学院	2015-10-22
郑光甫教授	2015-11-13
张志华教授 (上海交通大学)	2015-11-27
Dr. Peng Ding, Department of Statistics, UC Berkeley	2015-12-02
Dr. Linbo Wang, University of Washington	2015-12-03
Prof. Tony Cai, University of Pennsylvania	2015-12-07
Yoosoon Chang, Indiana University	2015-12-09
Dr. Theis Lange, Associate Professor in biostatistics, University of Copenhagen	2015-12-10
李凌云, 美国加州大学戴维斯分校	2015-12-13/15
姜祖恕, 台湾中央研究院	2015-12-18/23



## 五、论文发表

自 2015 年 1 月至 12 月，统计科学中心部分发表论文如下，其中 34 篇论文已经刊登，14 篇已经被相应期刊接收。

[1] An Y. and **Li C.** (2015). Approximating Local Volatility Functions of Stochastic Volatility Models: a Closed-form Expansion Approach, *Probability in the Engineering and Informational Sciences* 29, 547–563.

[2] Chang Jinyuan, **Chen Song Xi**, Chen Xiaohong (2015). High dimensional generalized empirical likelihood for moment restrictions with dependent data, *Journal of Econometrics*, 2015, 185, 1, 283.

[3] 陈家鼎, **李东风**(2015). 多元回归中选择自变量的一种简单方法, *应用概率统计*, 第 31 卷第 1 期, 71-88.

[4] Chen Zhen-Qing, **Ren Yanxia**, Song Renming and Zhang Rui (2015). Strong law of large numbers for supercritical superprocesses under second moment condition, *Front. Math. China*, 10(4)( 2015): 807–838.

[5] De Los Angeles, A., Ferrari, F., **Xi, R.**, Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H.G., Lensch, M. W., Lujan, E., Pei, D., Rossant, J., Wernig, M., Park, P.J., Daley, G.Q. (2015). Hallmarks of pluripotency, *Nature*, 525 (7570), 469-478.

[6] **He, Yangbo** and **Jia, Jinzhu** and Yu, Bin (2015). Counting and Exploring Sizes of Markov Equivalence Classes of Directed Acyclic Graphs, *Journal of Machine Learning Research*, 16, 2589-2609.

[7] **Jia Jinzhu** and Rohe Karl (2015). Preconditioning the Lasso for sign consistency, *Electronic Journal of Statistics* , pp 1150-1172.

[8] Jiang, Z. C., Ding, P. and **Geng, Z.** (2015). Qualitative evaluation of associations and validation of surrogacy by association transitivity, *Statistica Sinica*, 25, 1065-1079.

[9] Kuroda, M., **Geng, Z.** and Sakakihara, M. (2015). Improving the vector  $\epsilon$  acceleration for the EM algorithm using a re-starting procedure, *Comput Stat*, 30, 1051–1077.

[10] Lan, W., Luo, R., Tsai, C.L., **Wang, H.**, Yang, Y.(2015). Testing the Diagonality of a Large Covariance Matrix in a Regression Setting, *Journal of Business and Economics Statistics*, 33, 76-86.

- [11] Lee Tae-Hwy, **Tu Yundong** and Ullah Aman (2015). Forecasting Equity Premium: Global Historical Average versus Local Historical Average and Constraints, *Journal of Business and Economic Statistics*, 33:3, 393-402, July 2015.
- [12] Lee, L.F., and **Yu, J.** (2015). Estimation of Fixed Effects Panel Regression Models with Separable and Nonseparable Space-time Filters, *Journal of Econometrics* 184, 174-192.
- [13] Lee, L.F., and **Yu, J.** (2015). Spatial Panel Data Models. In: Baltagi, B (Eds.), Oxford Handbook of Panel Data, *Oxford University Press*.
- [14] **Li C.** and Li X. (2015). A Closed-form Expansion Approach for Pricing Discretely Monitored Variance Swaps, *Operations Research Letters* 43, 450–455.
- [15] **李辰旭**, 乔坤元, 高彬馨 (2015) 非线性动量交易策略: 理论和基于中国商品期货市场的应用, *金融学季刊*, 第9卷 第一期.
- [16] Li, Kang, Jiang, Bochuan, **Ai, Mingyao** (2015). Sliced space-filling designs with different levels of two-dimensional uniformity, *Journal of Statistical Planning and Inference* 157-158, 90-99.
- [17] Li, Kang, Zheng, Wei, **Ai, Mingyao** (2015). Optimal Designs for the Proportional Interference Model, *Annals of Statistics* 43(4), 1596-1616.
- [18] Liang, X., Zou T, Guo B., Li S., Zhang H., Zhang S., Huang H. and **Chen S. X.** (2015). Assessing Beijing's PM2.5 Pollution: Severity, Weather Impact, APEC and Winter Heating, *Proceedings of the Royal Society A*, 471, 20150257.
- [19] **Lin, W.**, Feng, R. and Li, H. (2015). Regularization methods for high-dimensional instrumental variables regression with an application to genetical genomics, *Journal of the American Statistical Association*, 110, 270–288.
- [20] Ma, Y., Lan, W., and **Wang, H.** (2015). Testing predictor significance with ultra high dimensional multivariate responses, *Computational Statistics and Data Analysis*, 83, 275-286.
- [21] Ma, Y., Lan, W., and **Wang, H.** (2015). A high dimensional two sample test under a low dimensional structure, *Journal of Multivariate Analysis*, 140, 162-170.
- [22] Nam, J.Y., Kim, N.K., Kim, S.C., Joung, J.G., **Xi, R.**, Lee, S. Park, P.J., Park, W.Y. (2015) Evaluation of somatic copy number estimation tools for whole-exome sequencing data, *Briefings In Bioinformatics*, bbv055.
- [23] Pan, R. and **Wang, H.** (2015). A Note on Testing Conditional Independence for Social Network Analysis, *Science China: Mathematics*, 58, 1179-1190.

- [24] Qian Junyang and **Jia Jinzhu** (2015). On Piecewise Pattern Recovery of Fused Lasso, *Computational Statistics and Data Analysis*, 221–237.
- [25] Qiu Yumou & **Chen Song Xi** (2015). Bandwidth Selection for High-Dimensional Covariance Matrix Estimation, *Journal of the American Statistical Association*, 110:511, 1160-1174, DOI: 10.1080/01621459.2014.950375.
- [26] **Ren Yanxia**, Song Renming and Zhang Rui (2015). Central limit theorems for supercritical superprocesses, *Stochastic Processes and Their Applications*, 125(2015) 428-457.
- [27] **Ren Yanxia**, Song Renming and Zhang Rui (2015). Limit theorems for some critical superprocesses, *Illinois Journal of Mathematics*, 59(1)(2015):235–276.
- [28] Su Liangjun, **Tu Yundong** and Ullah Aman (2015). Testing Additive Separability of Error Term in Nonparametric Structural Models, *Econometric Reviews*, 34:6-10, 1056-1087.
- [29] The Cancer Genome Atlas Research Network (Including **Xi, R.**) (2015). Comprehensive Molecular Characterization of Cutaneous Melanoma, *Cell*, 161(7):1681-96.
- [30] Wang Ruodu, Peng Liang, **Yang Jingping** (2015). CreditRisk+ Model with Dependent Risk Factors, *NAAJ* 19(1), 24-40.
- [31] Wang, W. and **Yu J.** (2015). Estimation of Spatial Panel Data Models with Time Varying Spatial Weights Matrices, *Economics Letters*, 128, 95-99.
- [32] Wu Yijun, Zheng Zhi, Zhou Shulin, **Yang Jingping** (2015). Dependence structure between LIBOR rates by copula method, *Frontiers of Mathematics in China*, 10(1): 147-183.
- [33] **Yang Jingping**, Chen Zhijin, Wang Fang, Wang Ruodu (2015). Composite Bernstein Copulas, *Astin Bulletin* 45(2), 445-475.
- [34] Zheng Yanting, Cui Wei, **Yang Jingping** (2015). Optimal reinsurance under distortion risk measures and expected value premium principle for reinsurer, *Journal of Systems Science & Complexity*, 28(1), 122-143.
- [35] Chen, H., Ding, P., **Geng, Z.** and Zhou, X. H. (2015). Semiparametric Inference of the Complier Average Causal Effect with Nonignorable Missing Outcomes, ( To appear in ACM Transactions on Intelligent Systems and Technology ).
- [36] **Chen, S.X.**, Lei, L.-H. and Tu, Y-D (2015). Functional Coefficient Moving Average Models with applications to forecasting Chinese CPI, ( To appear in Statistica Sinica ).



- [37] Guo, B. and **Chen S.X.** (2015). Tests for High Dimensional Generalized Linear Models, ( To appear in Journal of the Royal Statistical Society, Series B. ).
- [38] Jiang, Z. C., Ding, P. and **Geng, Z.** (2015). Principal causal effect identification and surrogate endpoint evaluation by multiple trials, ( To appear in J Royal Statist. Soc. B.).
- [39] Ju Dong, Qin Xin, **Xu Minya** and Direnzo S. Marco (2015). Boundary Conditoins of the Emotional Exhaustion-Unsafe Behavior Link: The Dark Side of Group Norms and Personal Control, ( To appear in Asia Pacific Journal of Management ).
- [40] Lee, L.F. and **Yu J.** (2015). Identification of spatial panel Durbin models, ( To appear in Journal of Applied Econometrics ).
- [41] Luo, P. and **Geng, Z.** (2015). Bounds on direct and indirect effects of treatment on a continuous endpoint, ( To appear in ACM Transactions on Intelligent Systems and Technology ).
- [42] Shi, J., Xing, Q., Duan, M., Wang, Z., Yang, L., Zhao, Y., Wang, Xiao, Liu, Y., Deng, M., Ding, Z., Ke, A., Zhou, J., Fan, J., Cao, Y., Wang, J.,\*, **Xi, R.\***, and Gao, Q.\* (2015). Inferring the progression of multifocal liver cancer from spatial and temporal genomic heterogeneity, ( To appear in Oncotarget ).
- [43] Wang, Yaping; **Ai, Mingyao**; Li, Kang (2015). Optimality of Pairwise Blocked Definitive Screening Designs, ( To appear in Annals of the Institute of Statistical Mathematics ).
- [44] **Xi, R.** and Lin, N. (2015). Direct Regression Modelling of High-order Moments in Big Data, ( To appear in Statistics and Its Interface ).
- [45] **Xi, R.**, Li, Y. and Hu, Y. (2015). Bayesian Quantile Regression Based on the Empirical Likelihood with spike and slab priors, ( To appear in Bayesian Analysis ).
- [46] **Xu, Minya**, Ping-Shou Zhong, and Wei Wang (2015). Detecting Variance Change-Points for Blocked Time Series and Dependent Panel Data, ( To appear in Journal of Business & Economic Statistics ).
- [47] **Zhang, J. L.**, Haerdle, W. K., Chen, C. Y. and Bommers, E. (2015). Distillation of News Flow into Analysis of Stock Reactions, ( To appear in Journal of Business and Economic Statistics ).
- [48] Zou, T. and **S. X. Chen** (2015). Enhancing Estimation for Interest Rate Diffusion Models with Bond Prices, ( To appear in Journal of Business and Economics Statistics ).

## 六、研究报告

- [1] **Mingyao Ai**, Xiangshun Kong and Kang Li  
A general theory for orthogonal array based latin hypercube designs
- [2] 北京大学统计科学中心, 光华管理学院  
空气质量评估报告
- [3] **Song Xi Chen**, Lihua Lei and **Yundong Tu**  
Functional coefficient moving average model with applications to forecasting Chinese CPI
- [4] Jinyuan Chang, **Song Xi Chen**, Xiaohong Chen  
High Dimensional Generalized Empirical Likelihood for Moment Restrictions with Dependent Data
- [5] **C. Li** and X. Li  
A Closed-form Expansion Approach for Pricing Discretely Monitored Variance Swaps
- [6] **C. Li** and Z. Zheng  
Evaluating Security Price Sensitivities Using Closed-form Asymptotic Expansions
- [7] **Yangbo He**, **Jinzhu Jia**, Bin Yu  
Counting and Exploring Sizes of Markov Equivalence Classes of Directed Acyclic Graphs
- [8] Zhang, X. and **J Yu**.  
Spatial Weights Matrix Selection and Model Averaging for Spatial Autoregressive Models.
- [9] Yang, Z, **J. Yu** and S.F. Liu.  
Bias Correction for Fixed Effects Spatial Panel Data Models.
- [10] Zhen-Qing Chen, **Yan-Xia Ren** and Ting Yang  
Boundary Harnack principle and gradient estimates for fractional Laplacian perturbed by non-local operators
- [11] Zhen-Qing Chen, **Yan-Xia Ren**, Renming Song and Rui Zhang  
Strong law of large numbers for supercritical superprocesses under second moment condition
- [12] Zhen-Qing Chen, **Yan-Xia Ren** and Ting Yang  
Law of large numbers for branching symmetric Hunt processes with measure-valued branching rates
- [13] Tao Zou and **Song Xi Chen**  
Enhancing Estimation for Interest Rate Diffusion Models with Bond Prices

- [14] **Yundong Tu**  
Efficient Estimation of Nonparametric Simultaneous Equations Models
- [15] Jing Zhou, **Yundong Tu**, Yuxin Chen, and **Hansheng Wang**  
Estimating Spatial Autocorrelation with Sampled Network Data
- [16] Jae-Yong Nam, Nayoung K. D. Kim, Sang Cheol Kim, Je-Gun Joung, **Ruibin Xi**, Semin Lee, Peter J. Park, Woong-Yang Park  
Evaluation of somatic copy number estimation tools for whole-exome sequencing data
- [17] **Ruibin Xi**, Nan Lin  
Direct Regression Modelling of High-order Moments in Big Data
- [18] Wei Zheng, **Mingyao Ai** and Kang Li  
Identification of universally optimal circular designs for the interference model
- [19] **Mingyao Ai** and Xiangshun Kong  
Central limit theorems for four new types of U-designs
- [20] Bochuan Jiang, **Mingyao Ai**  
Construction of uniform U designs
- [21] Xuan Liang, Tao Zou, Bin Guo, Shuo Li, Haozhe Zhang, Shuyi Zhang, Hui Huang, **Song Xi Chen**  
Assessing Beijing's PM<sub>2.5</sub> Pollution: Severity, Weather Impact, APEC and Winter Heating
- [22] Yaping Wang, **Mingyao Ai**  
Optimal designs for multiple treatments with unequal variances
- [23] Lu, X., Zhao, J., Chen, Y., and **Wang, H.**  
A Choice Model with a Diverging Choice Set for POI Data Analysis
- [24] Ma, Y., Lan, W., and **Wang, H.**  
A high dimensional two sample test under a low dimensional structure
- [25] Lan, W., Luo, R., Tsai, C.L., **Wang, H.**, Yang, Y.  
Testing the Diagonality of a Large Covariance Matrix in a Regression Setting
- [26] Pan, R., **Wang, H.**, and Li, R.  
Ultrahigh Dimensional Multi-Class Linear Discriminant Analysis by Pairwise Sure Independence Screening



[27] Yan, T., Qin, H., and **Wang, H.**

Asymptotics in undirected random graph models parameterized by the strengths of vertices

[28] Ma, Y., Lan, W., and **Wang, H.**

Testing predictor significance with ultra high dimensional multivariate responses

[29] Pan, R. and **Wang, H.**

A Note on Testing Conditional Independence for Social Network Analysis

[30] Huang, D., Yin, J., Shi, T., and **Wang, H.**

A statistical model for social network labeling

[31] Liangjun Su, **Yundogn Tu** and Aman Ullah

Testing Additive Separability of Error Term in Nonparametric Structural Models

[32] Jiyi Liu and **Jinzhu Jia**

Vanilla Lasso for sparse classification under single index models



Center for Statistical Science, Peking University

## 七、学术讲座

北京大学统计科学中心 2015 年共组织学术报告 49 场，报告人来自美国、英国、德国、澳大利亚、加拿大、新加坡、韩国、波兰、丹麦、香港和台湾以及国内各高校、研究机构和知名企业。其中来自海外的学者 36 场，国内的学者 13 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] Testing of High-Dimensional Covariance Structures

Yin Xia, UNC-Chapel Hill

2015-01-15 14:00-15:00

理科一号楼 1114 教室

### [2] Theory and computation for Bayesian nonparametric estimation problems

Dr. Shen Weining, The University of Texas MD Anderson Cancer Center

2015-02-24 10:30-11:30

理科一号楼 1418 教室

### [3] An additive structure for level sets of Gaussian fields

丁剑博士, 芝加哥大学

2015-03-09 15:00-16:00

理科一号楼 1303 教室

### [4] Discordance in estrogen receptor status between two primary breast cancers: impact of misclassification

Juxin Liu, Department of Mathematics and Statistics, University of Saskatchewan

2015-03-19 14:00-15:00

理科一号楼 1114 教室

### [5] Data mining and business analytics with big and small data

Johannes Ledolter, University of Iowa

2015-03-23 14:30-16:00

Room 217, Guanghua Building 2

### [6] Non-Nested Testing of Spatial Correlation

Peter M. Robinson, London School of Economics

2015-03-25 14:00-15:30

Room 216, Guanghua Building 2

### [7] Local Independence Feature Screening For Nonparametric And Semiparametric Models By Marginal Empirical Likelihood

Jinyuan Chang, The University of Melbourne

2015-03-31 14:00-15:30

Room 217, Guanghua Building 2

- [8] Shape-constrained Semiparametric Additive Stochastic Volatility Models  
Xinyi Xu, Ohio State University  
2015-04-02 14:30-15:30  
理科一号楼 1114 教室
- [9] Shrinkage Estimation of Common Breaks in Panel Data Models via Adaptive Group Fused Lasso  
Junhui Qian, Shanghai Jiao Tong University  
2015-04-09 14:00-15:30  
Room 217, Guanghua Building 2
- [10] Testing for Instrumental Variable Models under Weak Identification and Local Nonexogeneity  
Sung Yong Park, Chung-Ang University  
2015-04-16 14:00-15:30  
Room 217, Guanghua Building 2
- [11] Fluctuating-rate model and stochastic phenotype transition of a single cell  
葛颢研究员, 北京国际数学研究中心, 北京大学生物动态光学成像中心  
2015-04-20 15:00-16:00  
理科一号楼 1418 教室
- [12] Structural Change in Sparsity  
Myung Hwan Seo, Seoul National University  
2015-04-23 14:00-15:30  
Room 217, Guanghua Building 2
- [13] High-dimensional Statistical Process Control  
邹长亮, 南开大学  
2015-04-30 14:30-15:30  
理科一号楼 1114 教室
- [14] Simultaneous white noise models and optimal recovery of functional data  
Prof. Fang Yao, University of Toronto  
2015-05-12 14:00-15:00  
Room K02, 光华酒店一层
- [15] Gone with the Wind or Come With the Wind: the Spatial Externality of Air Pollution in China  
Mengmeng Guo, Southwestern University of Finance and Economics  
2015-05-21 14:00-15:30  
Room 217, Guanghua Building 2



- [16] An introduction to G-expectation  
宋永生副研究员, 中科院数学院  
2015-05-25 15:00-16:00  
理科一号楼 1479 教室
- [17] Robust Detection and Recovery of Structured Signals  
Tony Cai, University of Pennsylvania  
2015-05-26 14:00-15:00  
理科一号楼 1114 教室
- [18] Identification and Estimation of Hierarchy Effects in Social Interactions  
Zhongjian Lin, Emory University  
2015-05-27 14:00-15:30  
K1 of Guanghai Hotel
- [19] Sieve Maximum Likelihood Regression Analysis of Dependent Current Status Data  
胡涛, 首都师范大学数学科学学院  
2015-05-28 14:00-15:00  
理科一号楼 1114 教室
- [20] Cauchy semigroups: Nonlocally induced bound states  
Prof. Piotr Garbaczewski, University of Opole, Poland  
2015-06-01 15:00-16:00  
理科一号楼 1493 教室
- [21] Valid Post-Selection Inference  
Kai Zhang, UNC-Chapel Hill  
2015-06-03 14:00-15:00  
理科一号楼 1365 教室
- [22] Model Heterogeneity in Data Analysis: Detecting Clusters and Outliers via Cross-Validating Predictive Distributions  
George C. Tiao, University of Chicago  
2015-06-04 14:00-15:00  
Room 217, Guanghai Building 2
- [23] Survival Analysis of Loblolly Pine Trees With Spatially Correlated Random Effects  
Yili Hong, Virginia Tech  
2015-06-05 14:00-15:00  
理科一号楼 1364 教室
- [24] Propensity Score Adjustment for Estimating Population Causal Effect with Unequal Weights  
Bo Lu, Ohio State University

2015-06-11 14:00-15:00

理科一号楼 1365 教室

[25] Sieve Estimation of Pairwise Comparison Models

Jong-Myun Moon, University of College London

2015-06-16 14:00-15:30

K01 of Guanghua Hotel

[26] Bayesian Latent Propensity Score Approach for Average Causal Effect Estimation Allowing Covariate Measurement Error

Yi Huang, University of Maryland Baltimore County

2015-06-19 14:00-15:00

理科一号楼 1114 教室

[27] Discrete Fractal Dimensions and Large Scale Multifractals

肖益民教授, Michigan State University

2015-07-13 16:00-17:00

理科一号楼 1114 教室

[28] Latent Variable and Network Models for Measurement

Jingchen Liu, Associate Professor, Columbia University

2015-07-16 14:00-15:00

理科一号楼 1114 教室

[29] Approximate Bayesian Computation: treasure trove or trivial pursuit?

Simon Tavaré, Cancer Research UK Cambridge Institute & DAMTP, University of Cambridge

2015-08-11 14:00-15:00

理科一号楼 1114 教室

[30] Statistical Learning, Inference and Models for Big Data

Nancy Reid, Department of Statistical Science, University of Toronto

2015-08-12 14:00-15:00

理科一号楼 1114 教室

[31] Model Selection in Linear Measurement Error Models

台湾中央研究院, Chi-Lun Cheng

2015-09-17 14:00-15:00

理科一号楼 1479 教室

[32] Improving the robustness of doubly robust estimators

Peisong Han, University of Waterloo

2015-09-24 14:00-15:00

理科一号楼 1479 教室

[33] Semiparametric Estimation of Interval-valued Time Series Using Extreme Value Approach

Wei Lin, Capital University of Economics and Business

2015-09-24 15:00-16:30

Room 217, Guanghua Building 2

[34] A note on Constantin and Iyer's representation formula for the Navier-Stokes equations

罗德军副研究员, 中国科学院数学院

2015-09-28 15:00-16:00

理科一号楼 1479 教室

[35] Reconstructing intra-tumor phylogenies with Bayesian nonparametric models

Ke Yuan, University of Cambridge

2015-10-12 15:00-16:00

数学中心全斋全 29

[36] Innovated Interaction Screening for High-Dimensional Nonlinear Classification

郑泽敏, 中国科学技术大学

2015-10-13 14:00-15:00

理科一号楼 1479 教室

[37] Parsimonious, Accurate and Confident Credit Rating

Lijian Yang, Soochow University

2015-10-14 14:00-15:30

K01 of Guanghua Hotel

[38] Explosiveness and Initial Conditions: Theory and Application

Jun Yu, Singapore Management University

2015-10-15 14:00-15:30

K01 of Guanghua Hotel

[39] Estimation in Nonsmooth Moment Condition Models with Nonignorable Nonresponse

唐年胜教授, 云南大学数学与统计学院

2015-10-22 16:00-17:00

理科一号楼 1418 教室

[40] Some approximations results on the  $\Phi_3^4$  model

朱湘禅博士, 北京交通大学

2015-10-26 15:00-16:00

理科一号楼 1479 教室

[41] 醫學大數據的發展-台灣經驗

郑光甫 教授

2015-11-13 15:00-16:00



理科一号楼 1114 教室

[42] 机器学习中的非参数贝叶斯方法

张志华教授 (上海交通大学)

2015-11-27 15:00-16:00

理科一号楼 1114 教室

[43] Sensitivity Analysis without Assumptions

Dr. Peng Ding, Department of Statistics, UC Berkeley

2015-12-02 13:30-14:30

理科一号楼 1479 教室

[44] On Modeling and Estimation for the Relative Risk and Risk Difference

Dr. Linbo Wang, University of Washington

2015-12-03 10:30-11:30

理科一号楼 1418 教室

[45] Geometric Inference for General High-Dimensional Linear Inverse Problems

Prof. Tony Cai, University of Pennsylvania

2015-12-07 14:00-15:00

理科一号楼 1114 教室

[46] Econometric Analysis of Continuous Time Asset Pricing Models

Yoosoon Chang, Indiana University

2015-12-09 14:00-15:30

Room 217, Guanghua Building 2

[47] Recent developments within causal inference based effect separation (ie. mediation)

Dr. Theis Lange, Associate Professor in biostatistics, University of Copenhagen

2015-12-10 14:00-15:00

理科一号楼 1114 教室

[48] Central Limit Theorem for Linear Statistics of Eigenvalues of Large Random Matrices

李凌云博士, University of California, Davis

2015-12-14 15:00-16:00

理科一号楼 1479 教室

[49] Second order correction of European down and out options

姜祖恕教授, Institute of Mathematics, Academia Sinica

2015-12-22 15:00-16:00

理科一号楼 1418 教室

## 八、委员会

### 1. 国际顾问委员会



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。



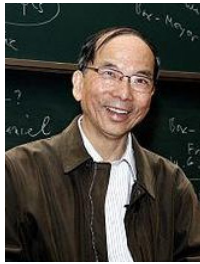
Peter G. Hall

委员，墨尔本大学和加州大学戴维斯分校的统计学教授，英国皇家学会院士，澳大利亚科学院院士，美国科学院外籍院士。



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、美国统计学会和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



陈大岳

委员，北京大学统计科学中心副主任，曾获国家级重大人才计划（基金委）等多项资助，现任北京大学数学科学学院副院长、中国数学会副理事长和《数学进展》主编等职。



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、流行病学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。



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”。



蔡洪滨

委员，北京大学光华管理学院院长，2006年获教育部新世纪优秀人才称号，2007年获评国家级重大人才计划（基金委），2009年被聘为讲座教授，2011年当选为世界计量经济学会会士(Fellow)，并于2012年被选为理事会理事。



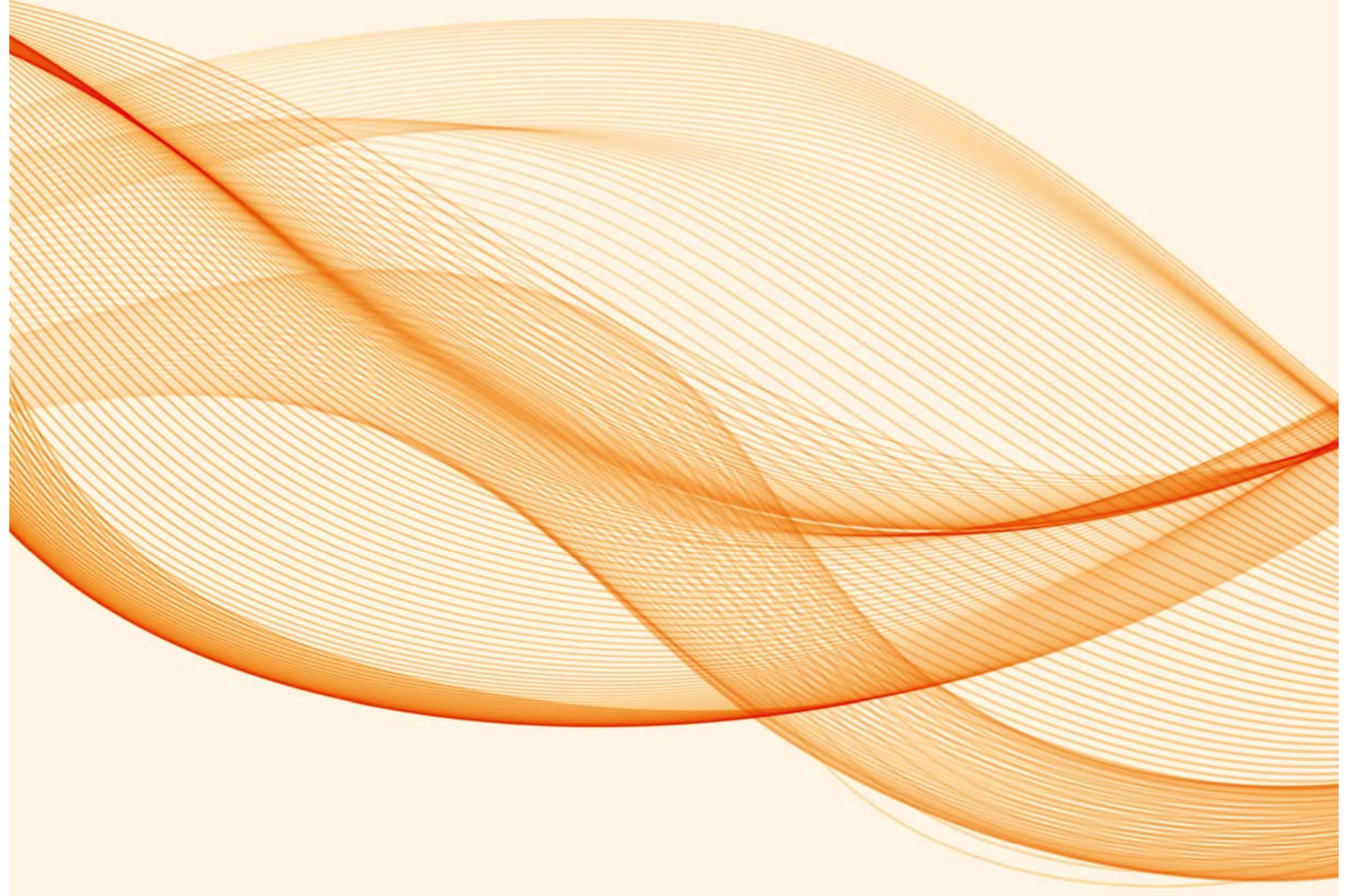
田刚

委员，北京大学数学科学学院院长,美国普林斯顿大学数学系教授，北京国际数学研究中心主任，中国科学院院士，全国政协常委，中国民主同盟中央副主席。在微分几何和数学物理领域做出了重大的贡献，曾经在2002年国际数学家大会上被邀请作大会报告。



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

An abstract graphic consisting of multiple overlapping, wavy lines in shades of orange and red, creating a sense of motion and depth. The lines are dense and form a mesh-like structure that flows across the page.

**Tel: 0086-010-62760736**

**Fax: 0086-010-62760736**

**Mail: [stat-center@pku.edu.cn](mailto:stat-center@pku.edu.cn)**

**Http: [www.stat-center.pku.cn](http://www.stat-center.pku.cn)**

**Add: 北京大学燕东园32号小楼**

**Postcode: 100871**