

大数据时代下的统计学^①

——第五届中国统计学年会综述

第五届中国统计学年会于 2014 年 10 月 18 日至 19 日在厦门大学隆重举行,来自中国、德国、新加坡等国家和地区的 300 余位代表和嘉宾莅临参会。本届年会主题为“大数据时代下的统计学”,以喻义统计学在新时代所面临的挑战和重大发展机遇。

本届年会特别邀请了国家统计局副局长许宪春博士、首都师范大学崔恒建教授、天津财经大学肖红叶教授、北京师范大学邱东教授、德国柏林洪堡大学 Wolfgang Härdle 教授、民生证券首席经济学家兼新华都商学院特聘教授邱晓华博士、中国人民大学金勇进教授、台湾辅仁大学谢邦昌教授以及厦门大学曾五一教授共 9 位学者做主题报告。报告内容涉及中国国民经济核算体系的修订问题、高维数据的部分惩罚似然比检验方法、社会经济大数据信息的最优采集与有效使用问题、经济测度的可加性问题、尾事件驱动的风险结构问题、中国经济分析与展望、企业创新调查抽样方案设计、云计算在大数据时代下统计研究发展的重要性以及统计学专业教学质量的国家标准等一系列问题。报告主题不仅从宏观层面涵盖了中国经济发展、大数据背景下的统计发展等内容,而且从理论和方法层面介绍了关于国民经济核算、指标质量、抽样方法、数据处理以及计算技术等方面的新变革。代表们对各位专家精彩的主题报告反应强烈,一致认为主题报告紧跟时代发展前沿,紧扣中国现实脉搏,反映了统计研究与工作的新发展。

本届年会共收到投稿论文近 400 篇,经组委会专家匿名评审,共录用 138 篇,作者包括高校教师、政府部门研究人员和在校学生等。总览论文质量,本次年会录用的文章写作规范,理论研究紧跟学科前沿,应用研究贴合中国现实,无论从深度还是广度上,都体现相当高的水准。

大会将入选论文分为 4 个小组进行分组报告交

流,并在每个分组报告会中设立了厦门大学—洪堡大学联合会场。本文按照选题内容和研究方法将论文分为以下 6 个专题对本次年会的主要学术观点进行归纳和综述。

一、大数据与统计学科发展

伴随着互联网电子商务的普及,智能移动设备的广泛使用,社交网络活动的日益频繁,以及云计算、物联网等技术的兴起,数据正在以前所未有的速度爆炸式增长,世界正在步入大数据时代。因此,如何运用和分析与日俱增的数据信息成为国内外学者热衷于思考和探索的问题。

厦门大学朱建平、李秋雅对大数据和大数据时代的概念进行了界定与辨析,剖析了现代信息技术和互联网技术的变革,并强调了高校必须重视利用大数据,改变高等教育领域的学习模式,提升个性化教育,辅助学生、教师的评价管理,促进科学研究的交流,为大数据时代下高等教育的发展道路提供了新思维。广东财经大学林洪、李棕杰提出统计学必须与计算机科学和信息科学整合发展以应对大数据时代的到来。湖南大学许涤龙、黄玉丹认为金融大数据的出现推动了传统金融模式的转变,一方面降低了统计成本、提高了统计时效并改善了统计效率,另一方面却增加了数据的安全隐患,甚至导致数据可衔接性差以及结构过于繁杂等问题。厦门大学童锦治、王晔等从大数据视角说明了税务大数据与传统税务信息化的差异、税务大数据对税务信息化发展的影响,以及现阶段税务部门建设税务大数据的思路和建议。

对于传统统计分析方法如何应对大数据到来所引发的数据规模、标准、形式等方面变化的问题,学者们从理论和应用两个方面给出了各自的观点。浙江工商大学程开明、徐梅芳从适用数据类型和非线

① 执笔人:厦门大学陈建宝、鞠芳煜、褚铸瑶。

性相关探测两个方面对传统相关分析进行了改进:一是拓展相关分析的适用数据类型,给出针对定类数据的相关分析方法;二是提出探测非线性相关的相关指数法和基于信息论基础的最大信息系数(MIC)方法。厦门大学刘维在交互作用分析框架下提出了一种通过压缩四维列联表变量以实现高维属性数据降维的方法,并通过实例验证了该方法具有较好的效果。天津财经大学杨贵军、孟杰提出了基于 AIC 信息准则的决策树模型评价和选择方法,实证结果表明,该方法可以有效替代和辅助错判率方法评价和选择决策树模型。上海财经大学葛守中和山东理工大学的朱志雯均运用大数据技术核算旅游活动数量,进而核算旅游排放的二氧化碳量,其中,葛守中以上海 2007 年的相关数据进行了实证研究,其旅游业并非低碳行业的结论引起了代表们的热烈讨论。

二、统计与计量的基础理论和方法

在模型构建、参数估计和假设检验等数理统计理论与方法的讨论过程中,重庆工商大学杨炜明提出了带有扩展得分函数的稳健方差估计量的修正二次推断函数,从理论上证明了修正二次推断函数更有效率,并通过数值模拟证明其在小样本情况下表现良好。厦门大学陈建宝在一个全新的框架下提出了具有随机效应的多元变系数空间滞后面板模型,构建了该模型的截面极大似然估计法,并证明了该估计量的一致性和渐近正态性, Monte Carlo 模拟结果表明该估计方法具有较好的小样本表现。北京大学宋晓军利用非参数残差经验分布和解释变量构造了检验条件分布对称性的检验统计量,采用自举法估计检验临界值, Monte Carlo 模拟结果表明该估计方法具有较好的小样本表现。兰州商学院刘明梳理了空间回归模型关于空间相关性检验、模型参数显著性检验、模型整体拟合优度检验以及模型选择检验等检验方法,并最终构建了空间回归模型的检验内容体系。天津财经大学李静华基于 PMM 多重插补法对线性模型的系数估计量进行了模拟研究,并得出在任意无回答机制下,随着插补重数增大,系数估计量的偏差和均方误差减小不显著的结论。厦门大学范新妍、方匡南等建立了半参数可加 β 回归模型及其估计量,模拟结果显示估计效果良好,并将该模型应用到家庭支出中医疗支出占比的案例中。

关于时间序列的长记忆性问题,德国汉诺威大学 Phillipp Sibbertsen 针对水平漂移和趋势缓慢变化会导致虚假长记忆的现象,基于多元局部 Whittle 似然函数提出了检验虚假多元长记忆的检验统计量,证明统计量不仅具有一致性,且能够显著提高检验功效。厦门大学李木易提出了动态混合双曲 GARCH 模型,讨论了其参数估计方法和渐近性质,与单变量长记忆模型对比,该模型具有更优的拟合效果和预测准确性。

在模型选择方面,浙江财经大学许林等考察了空间协方差结构误设时的模型选择问题,通过分别讨论工作协方差函数、工作方差函数和工作均值函数正确或误设的情形,评估了各种信息准则(如 AIC、BIC、RIC)寻找最优模型的能力,最终发现当样本量足够大时,BIC 和 RIC 准则的择优效果不受错误设定协方差的影响。

关于异常值和变点的探测方法,中国人民大学金立斌分别基于广义空间模型的均值扰动和方差扰动下的得分检验统计量提出了对应的异常值修正模型,并给出了相应的参数估计方法。上海财经大学李强对基于 LASSO 的带惩罚的变点检测研究方法进行了系统梳理和归纳,重点阐述了运用 LASSO 方法的时间序列模型、线性回归模型、分位数回归模型和面板数据模型等的变点检测问题,同时还探讨了该方法在不同领域的应用和前景展望,为国内学者对相关问题的研究提供了思路。山东工商学院袁靖等通过构建一致性估计量搜寻高维因子模型中因子载荷阵的不确定性,并持续估计前断或者后断因子的数量,最终得到通过增长变量个数模型的一致控制误差。

高维数据的变量选择方法是近些年来统计研究领域热点之一。厦门大学钟威提出了一种结合惩罚分位数回归和独立筛选程序的方法以识别单指数模型中超高维变量的重要程度,并通过 Monte Carlo 模拟考察了该方法的小样本表现。长春大学李华等在传统的马科维茨均值——方差最优化方法的基础上发展了谱修正估计方法,证明了谱修正估计量对于不同样本数、维数以及维数样本比具有稳定性等优良性质,模拟结果显示该方法克服了过度预测、过低预测等问题。厦门大学范青亮提出了工具变量回归中结构方程变量选择的自适应 LASSO 方法,并证明了估计量的优良性质, Monte Carlo 模拟结果表明

自适应 LASSO 方法相比于桥方法(Caner , 2009) 具有更小的偏差和均方差。暨南大学肖东莹、郑少智探讨了高维数据下半参数变系数可加风险模型的变量选择和估计, 并对美国诊所医疗数据做了实证分析。西南财经大学李扬研究了强分层面板数据中半参数模型的变量选择问题, 并提出了基于部分迭代估计方法的带有部分 MCP 惩罚项的似然函数, 利用 BIC 准则选择变量, 模拟结果表明该方法能够一致地选择最优模型且在有限样本下预测效果良好, 将其应用于中国股票市场数据的分析结果表明该方法能够有效选择变量。关于抽样理论方法的探讨, 西南财经大学夏怡凡等研究了配额随机抽样的理论性质, 并得到以下结论: 一是相关配额随机抽样与分层随机抽样一致; 二是在独立配额抽样前提下, 样本比例保持与总体比例一致, 样本均值估计量是总体均值的无偏估计; 三是对独立随机配额抽样、简单随机抽样和分层抽样的模拟检验结果显示, 独立随机配额抽样常常劣于分层抽样, 部分情况下优于简单随机抽样。

多元统计分析中一个重要的问题就是估计协方差矩阵, 中南大学谌自奇考察了一组非参数协方差模型, 运用局部线性方法估计经过 Cholesky 分解后的修正协方差矩阵, 估计量具有较好的理论性质, 并通过波士顿房地产数据阐述了该方法的有限样本表现。德国洪堡大学 Sebastian Holtz 研究了多维连续半鞅有噪声样本中潜在的积分协方差矩阵函数估计问题, 证明了其估计量的渐近正态性, 得到了估计量的最优速率和渐近协方差矩阵, 为解决此类问题的估计提供了有效方法。

另外, 广东财经大学林海明、陈军才为克服传统点图法降维图与原始数据实际不相吻合的缺点, 在对应分析优化模型框架下采用具有优良性的因子分析模型, 将其应用到美国 25 所高校办学问题的实例分析中。浙江工商大学李会杰研究了二维尺寸依赖更新风险模型, 得出两列独立同分布的非负向量在向量中元素非独立情形下具有精确的大偏差性质。中国人民大学李扬、孔雯婧提出了一种非结构化数据处理技术——基于主题的阈值调整半监督学习模型, 用于识别用户评论的情感倾向, 并对酒店评论数据做了实证研究。

三、国民经济核算与政府统计

近年来, 国民经济核算的热点问题主要涉及知识产权估算、国民财富衡量标准、GDP 国际比较、社会核算矩阵的设计和统计数据质量等内容。新修订的《2008 国民账户体系》决定将知识产权产品计入 GDP, 北京师范大学徐丽笑等对中国知识产权产品进行了估计, 结果显示过去 10 年间知识产权产品对 GDP 的平均贡献率为 2.23%, 相比发达国家产出仍处于较低水平。厦门大学的罗乐勤提出了新国民财富核算的概念, 将自然资源、人力资本和社会资本及管理纳入其中, 并对中国 1995—2009 年的新国民财富进行了重新估算。东北财经大学马晓君、任雅琪和国家统计局统计科学研究所冯蕾分别就美国 GDP 核算方法调整和欧洲账户体系修订对宏观经济数据的影响做了分析, 提出相关部门必须根据我国经济现实和统计能力有选择地对 SNA2008 的相关内容进行修订, 保持数据质量和可比性。浙江工商大学李佩瑾、徐蔼婷就 SNA2008 对中央银行产出核算的修订内容, 与 SNA1993 进行比较, 提出金融中介服务应从货币政策司的其他服务中独立出来实现单独核算, 成立专司金融中介服务的中国人民银行金融中介服务中心是改革我国中央银行产出核算的关键所在。北京师范大学席玮等参照 SNA2008 规定, 运用 PIM 方法对我国在 1998—2012 年的省级 R&D 资本存量做了重新估算。中国人民银行广州分行王开科及厦门大学曾五一针对我国统计数据质量问题展开讨论, 提出应把握事前控制、事中控制和事后控制这三个相互融合、循环往复的阶段以提高我国宏观经济统计数据质量, 并从“现有数据资料为基础的质量控制定量研究”和“前瞻性的数据质量控制定性研究”两方面探讨了提高我国宏观经济统计数据质量的思路。

鉴于我国重抽样设计、轻抽样估计的研究现状, 暨南大学陈光慧、刘建平拓展了超总体模型, 建立了一套广义回归系统, 并对中国农产量问题做了实例研究。山西财经大学马克卫、李宝瑜在中国社会核算矩阵交易流量表基础上, 依照经济循环过程对其进行了模块划分, 针对不同的模块分别设计了专用的流量转移模型, 实现交易流量向部门流量的平衡转化, 最终构造了一种崭新的、反映部门间经济流量变动情况的社会核算矩阵。浙江工商大学刘波、李

金昌设计了结合中国国情的非正规部门社会核算矩阵(FISAM),并使用相关数据和结果测算了正规、非正规和农业三个部门的中国FISAM。

国民经济核算和政府统计工作离不开统计指标的设计,江西财经大学郭露对传统特征价格模型进行优化,使用潜变量结构方程选择特征变量。杭州电子科技大学付荣将国民经济核算中自有住房的租金等价法与CPI中租金等价法做了比较研究,通过对比美国的处理方法,提出我国应遵循从易到难的规则,逐步考虑向美国规范的等价租金法靠近。中央财经大学胡永宏运用综合评价法构建了互联网可持续发展指数,利用2001—2011年反应全国互联网发展的10个指标,全面、系统地评价与测度了我国互联网的可持续发展水平,结果表明中国的互联网可持续发展总指数呈现逐年上升趋势,且其增长率呈现递减趋势,我国的互联网发展正逐步趋于平稳。上海交通大学朱启贵系统阐述了如何完善经济社会发展考核评价体系的问题,提出要立足中国国情,坚持“五位一体”,以引导政府与社会,促进经济社会协调发展。东北财经大学王岩从国民经济核算理论角度重新审视了使用购买力平价法所进行的国际比较研究,结论是国民经济核算体系(SNA)为基于购买力平价法的国际比较提供了最为基本的比较框架和方法论基础,基于购买力平价方法的国际比较可以看作是对国民经济核算体系的一种逻辑拓展,且已经形成两套较为系统的比较方法体系——基于支出法的ICP和基于生产法的ICOP。闽南师范大学陈立双参考相关国际标准和实践经验,针对中国CPI的低层汇总公式偏误、权重更新频率、商品质量变化、新产品偏误以及季节性产品处理等价格指数的系统性编制问题进行详尽的分析。东北财经大学谢长通过对分析2011年中国的价格水平、实际GDP和人均实际GDP等在世界上的总体情况,从发展中国家在国际比较项目(ICP)中处于不平等地位、ICP的合理性与局限性等方面说明应以客观、全面的视角看待ICP。北京师范大学吕光明、于学霖检验了我国食品价格上涨“二轮效应”影响,发现我国食品价格存在较为显著的“二轮效应”,通过提高通货膨胀预期和拉高工资水平渠道推高非食品和总体价格水平,且工资渠道的“二轮效应”要大于预期水平。

在加强统计数据数量收集的同时还应注重提升

数据质量。北京师范大学贾玮介绍了UNECE成员国利用超级连锁店扫描数据有效推进CPI编制的实践应用,提出中国相关统计部门应积极借鉴成功经验,拓宽CPI原始数据来源渠道和公式选择等建议。石河子大学党玮通过构建中国GDP与工业用电量、铁路运输量以及银行中长期贷款之间的向量自回归模型,分析了中国1985—2012年的GDP数据质量,结果表明我国GDP数据误差总体上在允许的5%范围以内,且数据质量有逐步改善的趋势。北京师范大学贾帅帅从对历史数据修订的考察、与国际收支统计及资金流量核算的衔接以及与外部数据的比较等角度评估了中国的GNI数据质量,结果表明中国GNI数据质量逐步改善、国民总收入核算与修订后的国际收支平衡表及资金流量表相互协调,说明中国国民总收入核算的水平有了长足进步。国家统计局统计科学研究所陶然从普查调查框误差测量的理论研究角度出发,研究了单位底册质量审核的理论方法。郑州华信学院李少利考察了中国国家与地方生产总值数据的修正发布制度,提出了“修正系数”的概念,从国家和地方两个层面,测度了1993—2008年我国经济普查对年度生产总值的修正效应。

关于投入产出模型及应用问题,福州大学郑珍远、郑姗姗等运用投入产出方法,构建了选择主导产业的指标评价体系,并实证分析了福建省的主导产业选择。厦门大学杨灿、郑正喜探讨了常规不加权和相应加权两种测度,并分别在Leontief和Ghosh模型体系下,测度需求拉动和供给推动的后向和前向关联效应,实证结果表明考虑了规模因素的加权测度方法更为客观。天津财经大学孙宪华采用投入产出法分析了中国1987—2010年我国制造业与生产性服务业各子行业的状况,探讨了生产性服务业与制造业的互动关联效应。中南财经政法大学吴文娟通过编制1990—2010年投入产出表序列,测算了投入产出表序列影响力系数、感应度系数等,结果表明我国金融业具有低耗能、高附加值的特征,与制造业关联密切,且金融业的发展受政策影响程度较大。

四、经济问题统计与计量

经济增长问题一直以来是国内外学者十分关心的问题之一,天津财经大学周国富等运用空间面板模型研究了中国经济增长的收敛速度,通过消除多

重共线性并考虑空间相关性等问题,得到中国经济增长的真实收敛速度为2.36%,并与国内外研究比较说明了该结果与我国经济的实际情况更相符。厦门大学褚铸瑶采用马尔科夫区制转移向量误差修正模型研究了中国政府消费与经济增长的短期波动模式与长期均衡关系,结果表明两者在样本区间内呈现显著的三区制特征。广州农商银行席晶等采用包含习惯形成特征的经济增长模型,分析了不同要素替代弹性下储蓄率与经济增长的变化关系。湖北大学危黎黎和向书坚基于季节时变平滑转换自回归模型,对我国工业增加值季度增长率的季节波动做了实证研究,结果表明我国工业增加值具有线性周期性波动和非线性季节波动的特征。江西财经大学徐斌使用非线性STR模型研究了政府消费与经济增长的关系,实证研究显示1955—1987年中国经济增长与政府消费在线性与非线性之间频繁转换,而1988—2011年仅存在线性关系。西南财经大学黎春使用上市公司数据构建了企业财务指数,发现宏观经济景气状态和企业财务水平存在长期均衡关系,企业营运、偿债和盈利能力都会对宏观景气状态产生影响。国家统计局统计科学研究所周晶对1980—2009年中国工业36个大行业CES生产函数中资本、劳动力、能源之间的三种嵌套形式做了非线性估计,结果表明多数行业拒绝C-D形式,说明CES函数嵌套较合适。在绿色经济研究方面,厦门大学钱争鸣、刘晓晨研究了绿色经济效率的测度,分别从微观和宏观层面探讨了提高绿色经济效率的理论机制,利用空间计量经济模型检验了理论结果,并分析了绿色经济效率的空间演化特征。

产业经济研究方面,山东工商学院王艳明、刘弈等通过构建耦合发展评判模型,发现山东半岛蓝色经济区海洋产业集聚系统与生态环境系统长期存在较为稳定的耦合交互作用,但耦合协调状态仍处于低水平。广州大学的张仁寿等采用VAR模型对广东省产业结构与经济增长的动态关系进行了实证研究,结果表明广东省已初步进入后工业化时代,产业结构尚存在偏差,部分产业发展质量不高。国家统计局统计科学研究所曹麦借鉴OECD《生产率测算手册》的研究方法,测算了2006—2011年间全国及各地区总量、产业及大类行业的全要素生产率(MFP),结果表明第二产业MFP最高,其中制造业MFP最高,40个大类行业MFP大于100%,陕西地

区MFP最高,河南MFP最低。由于国内对物联网产业统计指标体系的现有研究成果缺少针对性和完整性,为此,杭州电子科技大学薛洁构建了我国物联网产业统计指标体系的基本框架,提出了物联网产业统计指标体系的设计原则以及以增加值为核心指标的基本思路,为物联网产业的后续研究提供了参考和启示。兰州商学院王永瑜等应用投入产出分析方法研究甘肃省产业结构,得出甘肃省应优先发展文化服务业的结论。

为解答我国的货币政策规则是否有效,货币政策的实施是否对经济产生了预期影响等问题,河南大学牛永青构建了包含我国货币政策操作特征的DSGE模型,并得出在面对供给冲击和需求冲击时,货币政策的数量调控效果明显,而利率调控收效甚微的结论。安徽财经大学卢二坡基于频域格兰杰因果关系检验法对我国货币增长、通货膨胀与经济增长的关系进行研究,结果表明我国各层次的货币增长分别能在不同的频段为通货膨胀和经济增长提供预测信息,而各层次的货币增长在某些频段上是由通货膨胀或经济增长内生决定的。河北经贸大学彭国富构建了人民币内外价值调节模型,并实证验证了其可操作性。天津财经大学张肖基于DSGE视角,在封闭经济中构造了包含劳动市场雇佣摩擦的模型,对货币冲击和技术冲击的影响进行了考察。

在关于财政支出效率问题的研究上,国家统计局统计科学研究所何强在居民幸福最大化的假设下,构建了中央转移支付对地方财政支出影响的一般均衡模型,并利用1997—2012年中国31个省、市和自治区的面板数据进行了实证分析,得出中央转移支付对地方财政支出起到了引导作用且对地方财政主要支出类型的门槛效应均显著。厦门大学戴平生在结构分析法的框架下将地方财力均等化效应分解为反映中央财政干预的横向效应和反映地方财力聚集的纵向效应,得出中央的财政干预改善了我国地方财力的均等化。上海财经大学徐国祥考察了区域市场一体化对腐败的作用机制,指出其存在竞争效应和市场扩大效应。

在居民收入问题方面,北京师范大学的牛华辨析了财产性收入的内涵,分析研究了《中国统计年鉴》中的资金流量表和《城镇住户调查方案》中的部分项目。北京师范大学詹鹏、吴珊珊考察了遗产继承对财富分配的影响以及不同遗产税可能发挥的再

分配效应,研究发现遗产税的再分配效果较弱,将遗产税收入用于分配穷人则效果明显。中南财经政法大学向书坚、许芳研究了我国各省份 1988—2008 年产业结构对城乡收入差距的空间溢出效应和作用途径问题,结果表明我国各省份城乡收入差距和产业结构存在高度的空间相关性,并且两个因素都会产生显著的空间溢出效应。山西财经大学王涛将四分图模型引入收入差距测度中,直观而精确地反映了收入差距。浙江工商大学洪兴建提出了一种新的增值税收入分配效应分解方法,并运用 2007—2011 年我国城镇分组家庭的消费支出数据,分析了增值税“扩围”对居民收入分配的影响。

五、金融计量分析与风险管理

宏观经济走势与金融市场波动关系密不可分,湖南大学关家宜构造了马尔科夫贝叶斯向量自回归模型,研究了中国宏观经济对沪港两地股市的影响,并验证了贝叶斯方法的适用性。厦门大学赵华通过构建 VECM-ARJI-MGARCH 模型综合考察了中国股指期货和现货的协整关系、时变方差协方差特征以及常跳跃强度和时变跳跃强度下的期、现共跳性,实证结果表明,中国股指期货与现货市场具有显著的共跳性,跳跃强度表现出较高的持续性特征。浙江工商大学洪永婷基于 ICA-SV-t 模型的实证研究表明中国股票市场、债券市场和货币市场对我国外汇市场均存在波动溢出现象。华中师范大学王江涛提出了实现核估计的一类数据驱动最优窗宽选择算法,利用该算法研究金融市场高频数据,结果显示该算法不依赖初始值选取,具有良好的适应性和收敛速度。厦门大学陈海强、范云菲利用面板数据政策效应评估方法,研究了中国融资融券制度对股市波动率的影响,并得出融资融券制度的推出能够有效降低融资融券标的个股波动率的结论。

在投资需求和组合策略方面,天津财经大学李腊生将房产和股票引入居民资产组合,构造居民对资产品种的无差异曲线,推导出了股票需求函数。辽宁大学的隋云云采用区间数测度投资收益,利用半绝对偏差作为投资风险测度,在考虑到投资者的效用条件下,提出了一种新的基于效用的区间值投资组合模型。厦门大学沈雁通过协整方法选择配对交易资产组合,使配对交易不限于两个资产之间并扩展到多元配对交易,并用 LSTAR 模型基于均值回

复系数来确定阈值。

在风险度量方面,为改善信用风险模型的预测精度,台湾交通大学 Meng-Jou Lu 在巴塞尔新资本协定下,提出了利用动态单因子 Copula 模型估计日对日的相关系数。中南财经政法大学徐映梅基于 EVT-CoVaR 模型对系统重要性银行进行比较分析,发现系统重要性银行排名同期相对稳定,不同期呈动态变化,我国银行业系统性风险整体呈上升趋势。北京师范大学杜勇宏通过模拟不同宏观压力情景,对我国商业银行的信用风险压力进行测试,结果表明就均值而言风险可以控制,但在极端情形下风险偏高。中国人民大学王志刚给出了 Lee-Carter 模型完整的理论分布和区间预测表达式,通过与传统区间预测方法对比,发现传统方法倾向于得到较窄的预测区间,并低估长寿风险。厦门大学的吴红军对 2007—2013 年间我国 5 家主要信用评级公司的公司债评级精确度进行了测算,结果表明债券的等级高低与风险高低没有形成稳定一致的反向变动关系。暨南大学王斌会等将可拓理论用于地铁运营安全风险评价模型研究中,初步建立了基于多因素可拓理论的地铁运营安全风险评价和预警系统,并对广州市地铁运营安全风险进行实证分析。

此外,针对巨灾风险具有毁灭性强、发生概率低、分布呈现出严重的重尾形状的特点,浙江工商大学肖敏选取尾部较跳跃扩散过程更重的一个跳跃的 CEV 过程来模拟巨灾风险指数的路径,研究了巨灾债券的定价问题。新加坡管理大学的 Li Guo 对诸如比特币、alt 币等密码通货的潜力做了评价,并用计量方法研究其表现。

六、其他统计理论与应用

隐马尔科夫模型为探测基因组学中的拷贝数变异提供了有效的统计方法。山东大学王效强等考虑了一个同时包含多个修正隐过程的隐马尔科夫模型,克服了极大似然估计无法使用的问题,提出了一种基于变分法的推断算法来推断植物基因组的结构变异。

互联网在经济统计中扮演着举足轻重的角色。国家统计局统计科学研究所董倩以百度搜索指数为基础数据,运用多种模型对房产价格做了预测。天津财经大学袁铭在分布滞后模型的分析框架下,采用 Elastic net 方法进行收缩估计进而实现变量选

择,采用K重交叉验证技术确定最优惩罚因子和调整参数,建立了全国与城镇CPI的及时预测模型。中央财经大学李连友、屈增玉将互联网因素引入生产函数,运用分位数回归方法就互联网对经济发展的影响进行了定量分析,并用反事实分析方法对不同分位数下东、中、西部地区的经济发展水平差异进行了分解。

在环境生态问题研究方面,上海市统计局苑立波利用OECD行业数据库和联合国的COMTRADE数据库编制了包含中国、美国和世界其他国家(地区)的多地区投入产出表,并构建了中美对外贸易中的隐含碳排放量测算模型。中南财经政法大学温婷利用OECD新附加值贸易统计数据对中国2009年对外贸易隐含碳排放进行了重估算,并将其与传统贸易统计数据估算结果进行了比较分析。东北财经大学的王勇在投入产出结构分解分析框架下考察了中国的水足迹动因,结果表明:人均最终需求和人口是中国水资源永续性的主要动因;水资源利用效率、最终总需求和技术变化对水足迹的增长有负向影响等结论。北京师范大学的王亚菲利用全球多区域投入产出模型评估了经济全球化背景下给中国带来的环境压力,结果表明中国的环境压力是主要由

国外消费活动引起的,中国消费资源分布不均以及污染物排放等现象通过出口产品需求的增加加重了国内的环境压力。中国科学院、安徽财经大学崔连标等报告了后京都时代全球绿色气候基金的方案设计和影响。

在能源利用研究方面,北京师范大学赵楠在DEA-Bootstrap方法的框架下,对中国30个省级行政单位的能源利用效率进行了测算,研究发现Bootstrap技术确实有效纠正了传统DEA模型对能源利用效率测度值的偏差,中国各地区能源利用效率存在绝对收敛和相对收敛的特征,地区间差异在逐渐缩小,不同区域的能源利用效率收敛至不同的稳态,收敛速度也不尽相同。东南大学郭进研究了中国能源消费、贸易开放与经济增长的关系及区域差异,得出能源消费和贸易开放对中国经济增长存在显著的促进作用。东北财经大学李晓菲提出了基于经验模态分解方法和遗传算法以及一阶自适应系数模型的混合预测模型,利用山东半岛10分钟风速数据验证了该混合模型较传统模型有更好的预测精度。

(责任编辑:曹麦)

《统计研究》“统计图版式”要求

基本要求

1. 图的宽度大致以16开刊物的分栏宽度为准,长宽比例以美观适度为宜。
2. 图不取外边框。
3. 图内外文字或数字取6-8号为宜,宋体。
4. 图底色为白色。
5. 坐标刻度单位如为年份,请用“1997”,勿用“1997年”或“97”或“97年”;不得竖排或斜排。如年份较多,可以按间隔一年或数年的方式排出。
6. 坐标刻度单位如为变量值,一般不宜带有小数位,如“10000.0”;最多为5位数字,如需用6位数字标示,如“100000”,请将其变换为“10”,并改变相应的计量单位。图中标志值一般最多带两位小数,在特殊情况下,如整数位为0时,最多可带4位小数。
7. 图标题设在图外的图下方,小五号黑体居中,不得设在图内。
8. 图的标注放在图和图标题之间,以图外框左边界为准前空两格,小五号宋体。
9. 图的序号用“图1”表示,勿用“图1”、“图一”、“图1-1”等,文中只有1张图时,可用“图1”,亦可用“图”。