

# 交叉学科促进统计学的发展\*

朱建平 冯冲 梁振杰

**内容提要:** 进入20世纪,交叉学科不断涌现,其在认识和改造世界中的作用不断凸显。基于学科建设和科学研究的角度,学科的交叉可以分为4类,即学科门类内部的交叉、学科门类之间的交叉、学科门类与科学技术的交叉、自然学科与社会学科门类的交叉。在交叉学科的设置上,截至2021年6月,我国共有185所高校备案了616个交叉学科,其中应用经济学(经济统计学)和统计学涉及交叉学科的比例分别是13%和3.4%,统计学科在交叉学科的发展范围越来越广。在此基础上,本文探讨了统计学与其他学科交叉的内在动机和外部要求,并将2014—2021年中文社会科学引文索引和中国科学引文数据库中涉及的统计期刊发表论文作为研究对象,考察不同学科交叉程度的论文对统计学发展的作用力。最后,基于统计学未来可能的研究与应用动态,本文展示了统计学研究的10个重要发展方向。在交叉学科的促进下,各学科不断渗透、融合,构成了一个相互关联的多层级复杂网络。从学科交叉的角度研究统计学,将对社会进步和世界经济发展,特别是对促进智能化科学技术的飞速发展具有重要意义。

**关键词:** 交叉学科; 统计学; 未来发展

**DOI:** 10.19343/j.cnki.11-1302/c.2023.01.010

**中图分类号:** C816 **文献标识码:** A **文章编号:** 1002-4565(2023)01-0134-10

## The Cross-Disciplines' Promotion of the Development of Statistics

Zhu Jianping Feng Chong Liang Zhenjie

**Abstract:** As human society enters the 20th century, interdisciplinary disciplines have emerged continuously and play an important role in understanding and transforming the world. From the perspective of discipline construction and scientific research, the intersection of disciplines can be divided into four categories, namely, the intersection within disciplines, the intersection between disciplines, the intersection between disciplines and science and technology, and the intersection between natural disciplines and social disciplines. In the setting of cross-disciplines, as of June 2021, 185 universities in China have filed 616 cross-disciplines, among which the proportion of Applied Economics (Economic Statistics) and Statistics involved in cross-disciplines is 13% and 3.4% respectively, and the development of statistics discipline in cross-disciplines is getting wider and wider in scope. On this basis, this paper explores the intrinsic motives and external requirements of intersection between statistics and other disciplines, and takes the papers published in statistical journals involved in Chinese Social Science Citation Index and Chinese Science Citation Database from 2014 to 2021 as research objects to examine the influence of papers with different degrees of intersection of disciplines on the development of statistics. Finally, based on the possible future research and application trends of statistics, this paper shows ten important directions of statistical research. With the promotion of interdisciplinary disciplines, various disciplines are constantly infiltrating and

\*基金项目: 国家社会科学基金重大项目《大数据背景下重大传染性疾病传播路径及预警监测研究》(20&ZD137)。

merging to form a multi-level complex network. The study of statistics from a cross-disciplinary perspective is undoubtedly of great significance to the development of the world economy and social progress, and especially to the promotion of the rapid development of intelligent science and technology.

**Key words:** Cross-discipline; Statistics; Research Prospect

## 一、交叉学科的发展

20世纪以来，随着学科的发展与完善，学科纵向突破的难度急剧提升，这迫使科学研究聚焦于学科的横向转移与融合，也为交叉学科的发展提供了土壤（Larivière和Gingras，2010）。20世纪下半叶，交叉学科研究与应用成为热点。多学科的交叉与融合解决了许多学科内的难题，极大推动了科学的发展。在认识和改造世界的过程中，交叉学科的生命力和作用得到了充分证明（杜俊民，2020）。

国际跨学科协会（Association for Interdisciplinary Studies）于1980年成立后，已组织多次国际跨学科学术研讨会，以推动跨学科的科研与管理（程如烟，2005）。1984年，国务院颁布《关于科学工作的六条方针》，其中特别提到“自然科学中有与社会学交叉的学科”（李喜先，2001）。同年12月，中国社会科学院研究生院和中国科学院研究生院以“现代自然科学和社会科学”为主题，讨论了自然科学和社会科学的交叉与融合问题。1985年4月召开的全国首届交叉科学学术讨论会号召“迎接交叉科学的新时代”。近年来，有关交叉学科的文章数量越来越多。回顾历史，早在1986年《交叉科学》期刊创刊确立了“交叉学科”领域的开端；1987年《光明日报》出版社发布第一辑《交叉科学文库》，使得“交叉学科”的研究获得了较大发展（赵玉林，1995；郑晓瑛，2007）。

进入21世纪，中国科学院于2003年3月正式成立上海交叉学科研究中心。北京大学于2006年4月正式成立前沿交叉学科研究院，这也是我国高等院校首个跨学科研究试验田，此时交叉学科的理论研究水平有了明显提高。清华大学于2011年1月正式成立国内首个交叉信息科学研究机构，即交叉信息研究院，致力于世界顶尖交叉信息机构的建立与拔尖创新人才的培养。在此期间，各类科研机构 and 高等学校有针对性地设立研究交叉学科机构，同时，我国各省（区、市）科研机构与高校联合，根据研究特色纷纷构建交叉学科研究平台，对学科完善、科研创新和人才培养，以及推动社会全面进步具有十分重要的意义。

2020年11月，国家自然科学基金委员会成立交叉科学部，设有综合与战略规划处以及4个科学处，负责统筹国家自然科学基金交叉科学领域整体资助工作。2020年12月，国务院学位委员会、教育部颁布《关于设置“交叉学科”门类、“集成电路科学与工程”和“国家安全学”一级学科的通知》，按照《学位授予和人才培养学科目录设置与管理办法》的规定，经专家论证，国务院学位委员会批准，决定设置“交叉学科”门类（门类代码为“14”）。我国原有的13大学科门类正式升级为14大学科门类，分别为哲学、经济学、法学、教育学、文学、历史学、理学、工学、农学、医学、军事学、管理学、艺术学、交叉学科。

## 二、交叉学科的分类

纵观现代科学发展的轨迹，可以看到单学科研究模式已难以探讨复杂科学研究问题中的研究对象特征及其运动规律，学科间结构的变化已经展示出从学科分化演化为学科综合，从学科综合深化到学科交叉的趋势。就学科建设和科学研究的角度，交叉学科可分为4类，即学科门类内部的交叉、学科门类之间的交叉、学科门类与科学技术的交叉、自然学科与社会学科门类的交叉（朱建平，

1999)。

一是学科门类内部的有关学科相互交叉与渗透。随着学科纵向发展到一定程度,此时学科的专业和方向发展已较为完善,各专业内的突破创新难度显著增加,而学科内不同专业或不同方向之间的交叉与融合有利于学科整体水平的提升。例如,为解决信息、生命和环境等领域的核心基础科学问题,面向国际科学前沿和社会发展重大需求,数学、物理和化学等基础学科之间产生了物理与化学、数学与统计学、数学和力学等交叉。

二是学科门类之间的有关学科相互交叉与渗透。在科学研究和繁杂问题探讨中,人们发现学科门类之间有一种可以相互关联或协同使用的关系。例如,为探讨健康与生命相关的复杂系统作用机制,阐明生命现象中蕴含的科学原理,理学、工学、医学、农学等领域的研究不断交叉为人类健康与疾病防治作出了重要贡献。

三是学科门类与科学技术之间的交叉与渗透。加强学科建设是为了解决理论问题,强化技术是为了更好地解决实际问题,二者既有密切联系,又有重要区别。例如,当前大数据、人工智能和区块链等,就是建立在学科与技术之间的交叉。由社会发展中的重大需求推动,以解决工程或制造领域中的科学和技术瓶颈问题为导向,决定了学科门类与科学技术之间交叉的程度和方向。

四是自然科学与社会科学门类之间的交叉与渗透。伴随数学模型和计算机应用的普及,科学研究的群体化和社会化程度持续加深,自然科学家和社会科学家的研究对象相互交叉,知识在不同领域间渗透和融合。例如,解决人类可持续发展中的重大科学问题,人类经济、社会等复杂系统中的资源利用与规划问题,自然与人类社会的作用机制问题等都需要自然科学和社会科学的交叉与融合。

### 三、“统计学+”交叉学科的设置与建设

从学科设置角度来看,我国教育部学位管理与研究生教育司授予的“统计学+”交叉学科数量不断增加,其中,“统计学+”交叉学科定义为所属一级学科包含统计学的交叉学科。截至2019年5月31日,我国共授予153所高校(含508个交叉学科)自设交叉学科,其中“统计学+”交叉学科共16个。至2020年6月30日,我国新增8所高校(含53个交叉学科)自设交叉学科,其中“统计学+”交叉学科新增了1个,即大数据统计。至2021年6月30日,我国新增了24所高校(含80个交叉学科)自设交叉学科,其中“统计学+”交叉学科共新增4个,即生物信息学、车辆制造科学与工程、金融科技和大数据经济。

从交叉范围来看,“统计学+”交叉学科涉及范围愈加广泛,截至2021年6月30日,我国共有185所高校(含616个交叉学科)备案了交叉学科。“统计学+”交叉学科专业有21个(包含3所高校开设数据科学),分别为数据科学、能源经济与管理、工业与系统工程、经济信息管理、全球价值链、电子商务与信息管理等。我国交叉学科的一级学科分布结果显示<sup>①</sup>,至2021年,统计学的交叉学科比例在一级学科并列排名29,一级学科所占比例为3.4%。“统计学+”交叉学科涉及的学科范围越来越广,逐渐被医学、工学等学科所交叉。值得注意的是,一级学科应用经济学涉及交叉学科的比例达到13%,从专业设置的角度看,其中经济统计学专业在交叉学科建设中起到了举足轻重

<sup>①</sup>因篇幅所限,我国交叉学科的一级学科分布结果以附表展示,见《统计研究》网站所列附件。

的作用。

#### 四、“统计学+”交叉学科对统计学的作用力

科学体系的完善是一个螺旋上升的过程，总体经历由综合至分化，再至综合的演化路径。学科分化极大推动了现代科学体系的发展，形成了包括理学、哲学、经济学和管理学等13大学科门类。然而，值得注意的是，科学体系是一个多層级的系统，其内部的众多学科和理论是相互关联的，每个学科的发展和創新都受到其他学科的影响。而不同学科理论之间的非线性作用，即学科间理论的交叉和融合将更易产生新的理论，从而推动学科的发展，这也促使“交叉学科”门类成为我国第14个学科门类。同时，现实世界的复杂特性、社会需要和新技术的应用也要求学科之间的协同和交叉。

就统计学而言，统计学与其他学科交叉产生的影响将是双向的。从学科特性来看，统计学是一种横断科学，即以不同学科中具有普遍性的问题进行研究的学科。一方面，统计学在其他领域的应用将加强不同学科的联系和知识融合，并通过学科间的协同，应对各领域的复杂社会需求和挖掘研究客体的规律和特征；另一方面，这种交叉也将反过来促进统计学的发展。事实上，许多统计方法都是在解决特定应用问题时提出的，并且其他学科的发展，如数学和计算机学科也极大促进了统计学的发展速度。

“统计学+”交叉学科对统计学发展的影响可通过论文引用状况来衡量。现有研究对学科交叉程度与论文影响力二者关系探讨的结果并不一致，总体来看共呈现三种可能的结果，即学科交叉程度与论文的影响力高度正相关、负相关和呈倒U型关系（王菲菲等，2018）。为了验证“统计学+”交叉学科是否有助于统计学理论的发展，本文使用Porter和Rafols（2009）提出的综合性得分（Integration Score）测度统计学期刊论文的学科交叉程度，从而比较具有不同学科交叉程度的论文对统计学发展的作用力。

本文对学科交叉程度的度量如下：

$$I = 1 - \sum_{i,j} s_{ij} p_i p_j \quad (1)$$

其中， $p_i$ 为论文中引用文献属于学科*i*的比例， $s_{ij}$ 为学科*i*与学科*j*的相似系数。鉴于本文主要关注我国统计学的发展和学科交叉情况，在此所涉及的引用文献皆限定为中文文献。为了更准确地划分引用文献的学科类别，本文使用中国社会科学引文索引分类体系，将中文社会科学的学术期刊划分为25个大类。同时，为考察“统计学+”的学科交叉程度，本文将式（1）调整如下：

$$I = 1 - p \sum_j s_j p_j \quad (2)$$

其中， $p$ 为论文参考文献来自统计学期刊物的比例， $p_j$ 为论文参考文献来自学科*j*的比例。

另外，相对于Herfindahl或Shannon指数等交叉程度衡量指标，式（2）不仅能反映论文引用学科的集中程度，还能够以统计学与学科*j*的相似系数 $s_j$ 反映学科之间的差异。这意味着，当参考文献来源学科相似性很高时，式（2）的学科交叉程度接近于0。其中， $s_j = \min(\frac{N_j}{N}, 1)$ ， $N$ 为统计学期刊物中统计类参考文献的数量， $N_j$ 为统计学期刊物中属于学科*j*的参考文献数量。

本文将2014—2021年中国社会科学引文索引（CSSCI）和中国科学引文数据库（CSCD）中涉及的统计类期刊发表论文作为研究对象，论文“统计学+”的学科交叉程度分布如图1所示。本文共收集12539篇论文的引文和引证数据，其中参考文献74743篇，引证文献49423篇。从参考文献涉及学科来看，经济学、自然科学、管理学和地理学等学科文献数量最多，这表明统计学理论已在这些领

域得到广泛的传播和应用。从“统计学+”的学科交叉程度分布来看，约10%论文的学科交叉程度为0，说明其参考文献全部来自统计学刊物，而超过20%的论文交叉程度为1，即这些论文并未参考来自统计学刊物的论文。总体来看，其他论文的交叉程度呈阶梯上升状分布。

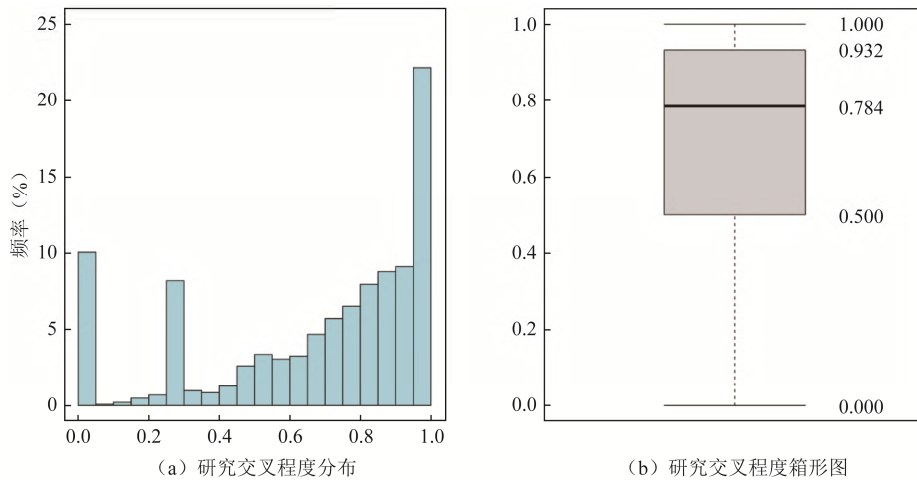


图1 “统计学+”的学科交叉程度分布

在对论文“统计学+”的学科交叉程度测度基础上，本文进一步探讨了具有不同交叉程度的论文对统计发展的作用力。如表1所示，本文分年份测算了“统计学+”的学科交叉程度四分位数。从历年交叉学科程度的下四分位数来看，统计学与其他学科的交叉程度总体呈递增态势，“统计学+”交叉学科的生命力得到充分证明。为考察不同学科交叉程度对统计学发展的作用力，本文根据“统计学+”的学科交叉程度上下四分位数将期刊论文分为三组，即高、中、低交叉组。继而，本文将论文在统计学刊物中被引用的次数作为其对统计学发展作用力的度量，并分年份衡量了不同学科交叉程度分组对统计学发展作用力的均值和标准差。

表1 “统计学+”交叉学科研究对统计学的作用力

		年份	2014	2015	2016	2017	2018	2019	2020	2021
		学科交叉程度	下四分位数	0.295	0.318	0.333	0.519	0.511	0.571	0.591
		上四分位数	0.949	0.943	0.946	0.933	0.936	0.939	0.92	0.916
统计学发展作用力	高度交叉	样本数	405	395	381	368	356	345	322	297
		均值	0.402	0.423	0.475	0.53	0.399	0.351	0.261	0.057
		标准差	0.878	1.011	1.045	1.192	0.828	0.801	0.724	0.247
	中度交叉	样本数	956	788	764	735	710	689	642	592
		均值	1.392	1.104	1.122	1.093	0.914	0.804	0.396	0.091
		标准差	3.981	1.955	2.204	2.373	1.422	2.368	0.954	0.342
	低度交叉	样本数	258	394	376	368	355	345	321	297
		均值	0.988	0.952	1.072	0.97	0.738	0.771	0.321	0.101
		标准差	1.895	2.085	2.228	1.918	1.326	1.593	0.711	0.446

总体来看，“统计学+”不同的学科交叉程度和交叉方向会对统计学发展产生不同的作用力。表1的结果显示，中度交叉组论文被引用次数相对更多，其次分别为低度交叉组和高度交叉组。这表明，适度的学科交叉将对统计学的发展产生更强的作用力。另外，各组对统计学发展的作用力均

值皆小于其标准差，由于作用力非负，由此可以推断在不同交叉程度分组中，少数交叉研究会统计学发展产生更大的作用力。学科间理论交叉和整合产生的作用并非是线性的，探索统计学与其他学科更有效的结合对统计学的发展意义重大。统计学产生于应用，基于复杂的社会需求和学科发展的内部动力，统计与其他学科统一整体的形成将成为统计学未来发展的趋势。

## 五、统计学的未来发展动态

从统计学的发展历史中可以看出，统计学是在应用过程中发展壮大，其生命力也体现于实践应用。目前，统计学在交叉学科的促进下已在以下领域获得广泛应用：第一，学科内部和学科之间的交叉，包括物理学和化学、生物数学和生物统计学、天体物理学和天文学、流行病统计和医疗统计、生态学和环境统计学、经济计量学、农业统计、人口统计、保险数学、心理计量学、法律和法医统计学、军事经济统计学等；第二，学科与技术的交叉，包括神经网络、区块链、人工智能、可靠性分析与寿命检验、数据库和可视化、网络媒介经营与管理、网络供应链管理、表决系统和政治学、管理科学与工程等。随着数据要素市场配置的不断优化，统计学的研究和应用体现出了新的能动性，也就是说，统计学的应用不断加深和创新的同时，也将会开辟出新的领域（朱建平，2014）。为了更明确统计学应用研究的动态及未来发展趋势，本文将从以下10个方面对统计学研究的重要领域及方向进行剖析。

### （一）处理社会经济复杂适应系统的统计学理论与方法研究

社会经济复杂适应系统是一个以人为核心，涉及人类各种活动复杂因素的庞大系统，其包括社会、经济、教育、科学技术及生态环境等领域，是一类重要的、典型的复杂适应系统（王晶华和施红玉，2002）。社会经济复杂适应系统具有若干特殊的情况与性质，使其难于描述、分析和控制。如主体和环境，即国家政府和国际组织，构成了系统的组元，当社会经济系统形成了一个有机的整体时，其主要原因是由主体与主体之间甚至包含主体与环境之间的相互作用所造成的。

社会经济复杂适应系统按照横向和纵向可以划分为许多相对独立的子系统。类似的例子还经常在环境、生态问题中遇到，如碳中和背景下我国经济高质量发展的系统性研究、重大传染性疾病的传播途径与监测监控等等。此外，在这类复杂系统研究与应用过程中，对于子系统的相互作用还需要深入探讨，会涉及到自然界中的诸多竞争关系，如物主与寄生物之间的竞争、种群生存之间的竞争或捕食系统的研究等等。

从统计角度研究社会经济复杂适应系统面临的挑战，主要体现在以下三个方面：第一，描述复杂适应系统需用成千上万个参数和变量，这是复杂高维数据研究的重要领域；第二，数据可获得性发生变化，打破了传统试验获得数据的方式，更多的数据往往只能通过观察得到，即难于保证数据的可重复性和随机性；第三，复杂系统将带来数据的复杂性，即数据来源、数据结构和数据格式等都呈现不均衡态势，给数据融合和数据整合提出更高的要求。这样，传统的统计学方法受到限制，需要研究全新的统计理论和方法进行处理。这一领域在国内外处于起步阶段，将会成为相当长时期内统计应用发展的重要趋势之一。

### （二）数据要素的开发与应用研究

近年来，随着大数据、互联网、物联网、云计算、人工智能等新兴技术的发展，人们进一步发现“数字”的经济重要性。如今数据要素作为经济发展新的生产要素，不仅给传统经济注入了新的动能，也加速推动全社会经济的“数字化”发展，促进与数据紧密相关的新模式和新业态的迅速崛起。那么，对数字经济产业进行科学核算，培育大数据市场，才能提高数据的可交易性，更好地推

动数据的开放共享；充分发挥数据要素新动能，开发数据新产品，加强应用服务，才能提升和实现数据资源的价值；进一步探讨统计法律的规范，清晰数据的产权交易，才能更好地实现数据的整合与保护。因此，这些问题给统计学提供了理论和应用的研究方向：第一，“数据”作为“要素”，其概念的理论界定和统计界定，统计对象的确定，包括数据与自然资源、资本、劳动和技术要素的联系与区别；第二，数据要素统计框架的构建，数据要素市场统计，数据要素生产收入核算；第三，数据要素的增长贡献及其对国民经济各部门、国家治理各领域的影响测度等。

### （三）数字经济运行机制的统计研究

数字经济是指以数据资源作为关键生产要素、以现代信息网络作为重要载体、以信息通信技术的有效使用作为效率提升和经济结构优化重要推动力的一系列经济活动。为贯彻落实党中央、国务院关于数字经济和信息化发展战略的重大决策部署，科学界定数字经济及其核心产业统计范围，全面统计数字经济发展规模、速度、结构，满足社会各界对数字经济的统计需求，国家统计局于2021年发布《数字经济及其核心产业统计分类（2021）》，确定数字经济产业范围为01数字产品制造业、02数字产品服务业、03数字技术应用业、04数字要素驱动业、05数字化效率提升业等5个大类（国家统计局，2021）。数字经济的发展给传统的经济统计带来了巨大挑战，如开展数字经济运行机制模拟及运行状况评价研究；数字经济核算的系统理论框架与方法体系研究；数字经济环境下宏观经济的统计监测及预警研究；数字经济与经济增长和其他产业的联动效应研究等。

### （四）大数据质量提升研究

随着智能产品的大规模普及，某些行业或企业通过互联网和物联网，人为产生大量的“虚数据”和“假数据”，这不是传统的统计数据质量问题，而且在数据类型发生巨大变化的同时，将这类数据以结构化数据、半结构化数据、非结构化数据形式融合在一个研究的统一体中。与传统数据相比，大数据时代下各种涉及数据运行的系统更容易产生数据质量问题，直接影响数据的存储、处理、分析等性能，给数据质量保障带来巨大的挑战，影响数据要素新动能的发挥（朱建平等，2019）。

在数据的生命周期中，影响数据质量的因素表现为信息、技术、流程和管理。为改进和提高大数据质量，确保数据功能，必须从大数据产生的源头抓起，从现代化管理入手，强化大数据质量管理的思想观念，把这一观念渗透到提升大数据质量的全过程，对数据运行机制进行全过程监控。从数据分析的角度，深入研究大数据形成过程中所遵循的客观规律，剖析其产生的内在机理，探索其科学有效的控制和改进措施。

### （五）网络产品的质量和生产率统计方法及应用研究

广泛意义下，网络产品是指在国民经济数字化基础上产生的物质产品、信息产品和网络服务等总称，包括通过数字化技术改造后的机器设备所生产出来的各种物质产品，在数字化平台上所生产流通的各类信息产品，以及通过数字化交易平台所提供的各式劳务。近年来，随着数字经济的发展，生产率和产品质量得到提高的互联网企业和网络产品获得了迅速的发展和提升。企业的管理者和决策者要提高网络产品的质量和生产率，分析网络产品生产全过程，那么科学的统计方法是必不可少的工具，这也是一个全新的研究领域，可以从统计角度探讨，在网络产品的过程控制、网络产品的可靠性等方面展开研究。第一，网络产品的统计过程控制研究迫切需要引进新思想，需要打破数据正态分布假定的条件，从根本上重新设计这些方法；第二，试验设计是一项重要技术，其核心目的是连续改进网络产品工程运行过程，保持数据采集过程的连续性，同时保证动态环境下的稳定性，因此探讨不同因素的水平设计将是根本；第三，针对网络产品寿命研究的模型构建还需做更多

的工作；第四，针对网络产品的生产效率研究，过去的一些工作表明，贝叶斯理论与方法将是分析和提高生产率的有效工具，同时动态环境的接受抽样也是该领域的重要技术之一。

#### （六）遥感信息与空间统计学的理论和应用研究

随着智能化推进遥感技术的不断发展，遥感数据的类型越来越丰富，数据的来源越来越多元化，遥感数据内在特性越来越鲜明，这给遥感大数据的智能化处理提出了新的挑战（宋维静等，2014）。空间统计学的大量应用，促使数据分析工作者对空间统计学理论和方法开展相关研究。第一，遥感大数据稀疏表征理论和方法的研究，基函数的构建和稀疏分解的方法研究；第二，时空数据挖掘技术的研究，典型例子包括重大传染病聚集性发现研究、古建筑或土地资源利用演变模式的发现研究、城市智能交通行为研究、飓风或地震等自然灾害发现研究；第三，遥感大数据内在结构特征和存在形式的建模问题，特别是针对机器学习机制，通过数据的降维、特征选择、模式识别和知识表达等技术，研究如何有效利用不同类型的遥感数据和信息。

#### （七）社会经济大数据计算研究（计算社会经济学）

随着智能化技术的飞跃发展，人类参与各种社会活动的行为数据，以及社会发展过程的历史数据，被越来越多的传感设备和数据采集终端记录下来。这些数据是通过互联网、物联网、智能遥感、移动通信、社交网络等产生的新型数据，而不是传统的社会科学处理的数据字段和惯常数据。这些数据的分析与理解，需要先进的机器学习和数据挖掘技术，这给社会科学研究者提出了新的要求。第一，这些数据来源广、精度高、规模大、实时性强，可以提高数据的稠密性，为多源数据整合开辟新的研究领域；第二，通过对新型数据的分析，可以更好地感知社会经济新态势，探讨和孕育新理论，发现可能的异常，预测未来发展趋势等等；第三，基于大规模真实数据对新型数据的研究，特别是对与社会发展过程有关的经济问题，以及与经济发展有关的社会问题研究，将开辟应用统计学研究的新领域。

这一新型数据的研究，充分体现了社会科学、经济学和计算机科学理念与方法的深度融合，可以将其界定为计算社会经济学，其应用能够扩展到司法领域中法庭第三方数据证据的分析与认定，使用物联网数据进行决策科学化的应用，以及使用通信数据进行金融风险防控方面的应用等等。

#### （八）健康医学和生物卫生中的大数据分析应用研究（计算社会医学）

随着大数据应用范围的不断扩展，健康医学、公共卫生和生物科技领域逐渐发展形成惠及全民“互联网+大健康”的新兴业态，这使得人民群众对美好健康生活期望值不断提高。在推动数据科学技术与健康医疗、公共卫生和生物科技深度融合的过程中，开展大数据应用既是机会同时也面临着巨大的挑战。健康医学和生物卫生大数据应用研究的方向是社会化医学和个性化医学，即以健康医疗大数据研究为基础，通过新的模式引导政务、教育、商业、环境等各个领域的变革；通过新的思维为个体量身设计出最佳治疗、康养、养老、社保等方案，以期达到最佳的定制健康医疗模式。该领域研究可以界定为计算社会医学，其所涉及大数据的开发环节，即存储数据、收集数据、挖掘数据、数据转化实用，以及智能化产品的实现，都离不开统计理论和方法的应用：第一，通过健康医疗大数据应用，促进优质医疗资源科学合理配置，充分体现医疗资源的均衡性，努力提高人民群众幸福感；第二，通过健康医疗大数据研究，深化医疗改革，为优化多级医疗服务体系，进一步实现异地结算和远程服务，注入新动力；第三，通过健康医疗大数据应用，开展健康科技新产品研发，创建健康服务新态势，推进上下游全健康产业链的发展。

#### （九）统计学方法在企业数字化转型中的应用研究

在微观层面，数字经济推动企业治理和互动模式的深刻变革，越来越多的企业开始重视和关注



数字化转型中统计学在企业现代管理中的应用。随着社会经济的不断发展与进步,现代企业逐渐加快了推动数字化转型的步伐,将组织创新、技术创新、跨界创新、融合创新,作为实现企业走向数字化发展和高质量发展的重要标志。数字化转型就是有目标地开发数字化技术,利用数据和数字技术的力量,支持并创建一个充满活力的数字化商业模式。而统计学理论支撑和应用,将直接关系到企业数字化转型的质量和效率,统计学可以为企业提供真实可靠的各项数据信息,帮助企业在数字化转型过程中,做出科学决策,创造出更多经济和社会效益,提升企业可持续发展的能力。有关的研究主要集中在以下几个方面:第一,运用统计学理论和方法,建立完善的企业数字化转型评价指标体系;第二,利用大数据、人工智能、区块链等创新技术,构建数字化转型机制,获取企业运营各环节的核心数据,建立企业的动态数据模型;第三,运用统计学方法,并结合行业大数据构建预警监测系统,加强企业核心竞争力;第四,将人工智能和现代统计学方法相结合,对企业生产经营过程实施科学的设计和控制在等等。

#### (十) 计算机软件中的统计理论研究

统计模拟与计算机软件是实现统计理论方法应用的重要支撑,也是充分体现统计理论方法赋能人工智能的重要工具。该领域的问题主要体现在两个方面:一是开发计算过程中,有效的统计计算理论和方法;二是开发软件过程中,通过计算机实现理论方法的应用。前者更依赖于统计学理论工作者,重点是理论和方法的研究;后者需要统计学与计算机相互配合才能解决,重点是算法和技术的研究。主要研究方向包括大数据集合与多源数据的处理方法、集成算法在统计计算中的应用、统计计算理论与数据挖掘技术比较研究、统计软件设计与开发及相关研究等。第一,数据挖掘技术中的统计学基础理论研究,以及通过统计理论方法的深入研究,提升数据挖掘技术水平;第二,大数据可视化技术中统计学内在含义的研究,实时实现交互图象数据分析;第三,行业或专业的统计软件开发研究,例如环境科学、流行病学、心理计量学、地质学、语言学、制图学及其他领域的应用软件。

本文所列的统计学未来研究重要领域并非是相互独立的,而是构成了一个相互关联的多层级复杂网络。伴随大数据时代的迅速发展,统计学将逐渐渗透到自然科学和社会科学领域,将与计算机、互联网、物联网、大数据、智能技术、区块链、自动化、云架构等理论和技术相结合,以解决自然界和自然发展过程中的疑难问题。如今越来越多的学者将尖端科学技术视为研究目标,这也为统计学应用提供了新的发展方向,进一步促进了统计学与尖端科学技术相互融合。而各门学科不断融合,促进现代科学整体化发展的趋势,并形成了一个相互联系的整体(朱建平和张悦涵,2016)。例如,本文数字经济运行机制的统计研究部分侧重于数字经济的宏观核算、调控和治理,而大数据质量提升研究、网络产品的质量和生产率统计方法及其应用研究、统计学方法在企业数字化转型中的应用研究分别聚焦于数字经济的不同维度和视角。事实上,在数字化进程中,数据要素的作用和地位不断凸显,对数字经济相关的测算和运行机制的探讨需在数据要素统计框架和数字治理的基础上进行,而这必然要求研究人员控制人为产生的“虚数据”和“假数据”。同时,各领域获取的海量数据和应用需求将进一步推动社会经济复杂适应系统的统计学理论、计算社会经济学和计算社会医学等领域的相关研究,而相应成果将反过来促进对数字经济领域的测算和机制探讨。不同领域的交互将是螺旋上升的。由此,统计学未来发展的研究,将对全社会进步和世界经济发展,特别对促进智能化科学技术的飞速发展具有重要意义。

## 参考文献

- [1] 程如烟. 美国国家科学院协会报告《促进跨学科研究》述评[J]. 中国软科学, 2005(10): 154-156.
- [2] 杜俊民. 试论学科与跨学科的统一[J]. 科学技术与辩证法, 2000(4): 56-59.
- [3] 国家统计局. 数字经济及其核心产业统计分类(2021) [EB/OL]. [http://www.gov.cn/gongbao/content/2021/content\\_5625996.htm](http://www.gov.cn/gongbao/content/2021/content_5625996.htm).
- [4] 李喜先. 论交叉科学[J]. 科学学研究, 2001(1): 22-27.
- [5] 宋维静, 刘鹏, 王力哲, 等. 遥感大数据的智能处理: 现状与挑战[J]. 工程研究-跨学科视野中的工程, 2014, 6(3): 259-265.
- [6] 王菲菲, 贾晨冉, 刘俊婉, 等. 交叉学科中文献学术价值的成长与老化研究——以医学信息学为例[J]. 科学学与科学技术管理, 2018, 39(2): 11-22.
- [7] 王晶华, 施红玉. 从系统科学角度看学科交叉现象[J]. 科学学与科学技术管理, 2002(12): 5-8.
- [8] 赵玉林. 交叉科学的崛起及其形成机制[J]. 自然杂志, 1995(3): 157-159.
- [9] 郑晓瑛. 交叉学科的重要性及其发展[J]. 北京大学学报(哲学社会科学版), 2007(3): 141-147.
- [10] 朱建平. 世纪之交中国统计学科的回顾与思考[M]. 中国经济出版社, 1999.
- [11] 朱建平, 谢邦昌, 马双鸽, 等. 大数据——统计理论、方法与应用[M]. 北京大学出版社, 2019.
- [12] 朱建平, 章贵军, 刘晓葳. 大数据时代下数据分析理念的辨析——基于统计思维的研究[J]. 统计研究, 2014, 31(2): 10-19.
- [13] 朱建平, 张悦涵. 大数据时代对传统统计学变革的思考[J]. 统计研究, 2016, 33(2): 3-9.
- [14] Larivière V, Gingras Y. On the Relationship Between Interdisciplinarity and Scientific Impact[J]. Journal of the American Society for Information Science and Technology, 2010, 61(1): 126-131.
- [15] Porter A, Rafols I. Is Science Becoming More Interdisciplinary? Measuring and Mapping Six Research Fields Over Time[J]. Scientometrics, 2009, 81(3): 719-745.

## 作者简介

朱建平, 厦门大学管理学院教授、博士生导师, 厦门大学健康医疗大数据国家研究院副院长, 厦门大学数据挖掘研究中心主任, 湖北经济学特聘讲座教授, 教育部高等学校统计学类专业教学指导委员会副主任。研究方向为数理统计、数据挖掘。

冯冲, 厦门理工学院数学与统计学院讲师, 厦门大学数据挖掘研究中心研究人员。研究方向为数据挖掘、数据科学与商业智能。

梁振杰(通讯作者), 闽江学院经济与管理学院讲师, 闽江学院产业经济与公共政策研究中心研究员, 厦门大学数据挖掘研究中心研究人员。研究方向为社会经济统计、数据挖掘。电子邮箱: [xmzjliang@163.com](mailto:xmzjliang@163.com)。

(责任编辑: 刘雅欣)