

【研究动态】

人大复印报刊资料《统计与精算》 2014—2019 年转载情况的文献计量分析

张 皓 车明佳

【摘 要】以复印报刊资料《统计与精算》2014—2019 年全文转载文章为研究对象,通过对转载的高质量论文的全文进行词频分析,探究统计与精算学科近 6 年总体的研究热点;同时利用论文的标题、摘要、关键词等的共词信息,使用文献计量学和 Citespace 5.7R1 软件进行科学知识图谱的可视化,深入分析统计与精算学科的知识结构与研究趋势的变化,并结合统计与精算学科论文作者的科研合作网络,揭示此领域内学术合作关系的规律。结果表明,该刊转载的文章以问题为导向,侧重研究问题的时代性和前沿性;高校和研究机构是统计与精算学科发展的重镇;学科带头人、专家团队对所在机构的学科建设起引领作用;专业期刊是统计与精算学科学术成果传播的主阵地;发表论文篇数较多的机构均选择与其他机构进行合作;多作者成为统计与精算领域学术成果发表的显著特征,学者们通过优势互补、资源共享实现共赢;复印报刊资料《统计与精算》通过博观约取、优中选优,为推动学科发展和学术传播做出了自己独特的贡献。

【关键词】人大复印报刊资料;统计与精算;论文转载;文献计量;知识图谱

【作者简介】张皓,女,北京人,中国人民大学书报资料中心副编审,研究方向:统计学、市场营销学;车明佳,女,山西太原人,中国人民大学统计学学院博士生,研究方向:统计学(北京 100872)。

【原文出处】《统计与信息论坛》(西安),2020.12.115~123

一、引言

统计学自改革开放以来取得了快速发展,特别是在晋升为一级学科以后,更是成为了显学,不仅有近百所学校开设了相关专业,统计方法也在科学研究中得到了广泛的应用。作为新中国最早从事人文社会科学文献搜集、整理、编辑、出版的学术信息资料提供和服务机构,中国人民大学书报资料中心编辑出版的复印报刊资料系列刊在中国人文社会科学学术研究和高等教育繁荣发展的进程中做出了特殊贡献。复印报刊资料《统计与精算》于 1984 年创办,最初的名称是《统计学、经济数学方法》,2001 年改为现刊名,沿用至今。该刊定位在应用统计方向,侧重社会经济统计,意图在办刊中体现中国人民大学的特色——精选有关国民经济核算、收入分配统计、人口统计、产业统计、生物与医疗卫生统计、环境与生态统计等领域的文章,兼收

风险管理与保险精算的内容,秉持“学术为本,为教学科研服务”的办刊宗旨及“精选千家报刊,荟萃中华学术”的编辑方针,忠实记录中国应用统计学的沿革与发展。复印报刊资料《统计与精算》以研究内容是否具有创新性、论证是否完备、对社会发展进步产生的推动作用的大小、论题的复杂程度和资料收集处理的难易程度为遴选依据,始终保持了较高的学术水准及较广泛的社会影响力,日益受到学术期刊界、教学科研机构和广大人文社会科学工作者的广泛关注和高度重视,成为各类学术评价的重要依据之一,对促进中国统计学学术期刊发展和学术传播、推动学术繁荣发挥了重要作用^[1-3]。

随着移动互联网的发展和数据科学的成熟,以知识图谱为代表的人工智能技术,因为拥有强大的显示科学知识的发展进程与结构关系的功能而被广泛使用,特别是用来观察某个领域的研究趋势或

动向,并以可视化的方式呈现。为了展示统计与精算学科的演进路径,本文通过对《统计与精算》杂志近6年转载情况的分析,呈现该刊的特点和该领域的研究热点、重要学者和研究机构,回应各界对复印报刊资料的关切。

二、数据来源和研究方法

本文所采用的分析数据来源为复印报刊资料《统计与精算》2014-2019年所收录的372篇全文转载文章的文本数据(剔除1篇从报纸上转载的文章),包括标题、摘要、关键词、作者信息、正文内容等。通过对转载的高质量论文的全文进行词频分析,探究统计与精算学科近6年总体的研究热点;同时利用论文的标题、摘要、关键词等的共词信息,使用文献计量学 Citespace 5.7R1 软件进行科学知识图谱的可视化,深入分析统计与精算学科的知识结构与研究趋势的变化;并结合统计与精算学科论文作者的科研合作网络,揭示此领域内学术合作关系的规律。

三、数据分析

(一) 文章类型与来源

《统计与精算》2014-2019年共转载文章372篇,索引文章9448篇。其中,2014年转载文章60篇,索引文章1861篇;2015年转载文章67篇,索引文章1660篇;2016年转载文章69篇,索引文章1361篇;2017年转载文章61篇,索引文章1475篇;2018年转载文章58篇,索引文章1559篇;2019年转载文章57篇,索引文章1532篇。年均转载文章62篇,索引文章1575篇,转载比例3.95%。

从转载的文章的内容分类看(见图1),有关风险管理与精算的文章74篇,占19.89%;有关经济统计的文章64篇,占17.20%;有关国民经济核算的文章62篇,占16.67%;有关人口统计的文章47篇,占12.63%;有关理论与方法的文章39篇,占10.48%;有关公共部门统计的文章36篇,占9.68%;有关社会发展统计的文章18篇,占4.84%;有关环境与生态统计的文章17篇,占4.57%;有关学科建设的文章15篇,占4.03%。

从转载文章的长度来看(见图2),2014年转载文章的平均长度为15713字,2015、2016年较为稳

定,但2017-2019年的平均长度分别为17754、18600、17963字,虽然原发刊文章的平均长度提升,但长篇幅文章字数呈缩减的趋势。

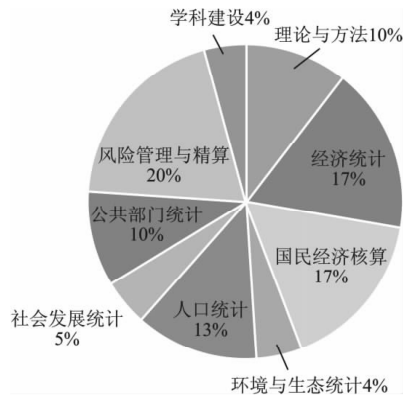


图1 《统计与精算》2014-2019年转载文章分类图

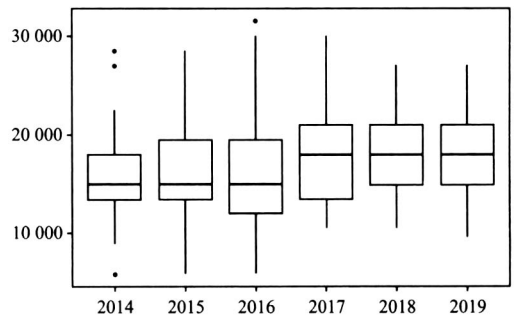


图2 《统计与精算》2014-2019转载文章字数图

从转载文章的来源看(见表1),2014-2019年《统计研究》共转载文章64篇,占17.20%;《数量经济技术经济研究》转载51篇,占13.71%;《数理统计与管理》转载25篇,占6.72%;《保险研究》转载20篇,占5.38%;《统计与信息论坛》转载19篇,占5.11%;《经济学动态》转载10篇,占2.69%;《经济研究》《中国农村经济》各转载9篇,分别占2.42%;《农业技术经济》转载7篇,占1.88%;《统计与决策》《中国工业经济》《人口与经济》各转载6篇,分别占1.61%;《经济学》《管理世界》各转载5篇,分别占1.61%;《经济统计学》《劳动经济研究》《中国管理科学》《调研世界》《山西财经大学学报》分别各转载4篇,各占1.08%。转载4篇以上的刊物共计19种,共转载文章262篇,占全部转载文章的70.43%。其中,6本统计类专业期刊(含集刊)共转载122篇,相关学科期刊共转载71篇,二者所构成的泛统计类期刊占比为51.88%;非统计类期刊为

表1 《统计与精算》2014—2019年转载文章来源

期刊名称	转载篇数	占比(%)	排序
专业期刊			
统计研究	64	17.2	1
数理统计与管理	25	6.72	3
统计与信息论坛	19	5.11	5
统计与决策	6	1.61	10
经济统计学	4	1.08	15
调研世界	4	1.08	15
非专业期刊			
数量经济技术经济研究	51	13.71	2
保险研究	20	5.38	4
经济学动态	10	2.69	6
经济研究	9	2.42	7
中国农村经济	9	2.42	7
农业技术经济	7	1.88	9
中国工业经济	6	1.61	10
人口与经济	6	1.61	10
管理世界	5	1.34	13
经济学	5	1.34	13
劳动经济研究	4	1.08	15
山西财经大学学报	4	1.08	15
中国管理科学	4	1.08	15
合计	262	70.43	

179篇,占比48.12%。

从以上的数据可以看出,第一,作为一本定位于应用统计的学术期刊,《统计与精算》杂志收录的文章主要是有关经济统计、国民经济核算、人口统计和风险管理与精算方面的内容,纯理论的文章较少,且以研究动态为主。第二,来源刊发文长度呈增加趋势,较短和较长篇幅文章均向中篇幅文章靠拢。第三,该刊转载的文章,专业刊和非专业刊比例大致相当。这是因为,首先,统计类专业期刊如《统计研究》《统计与信息论坛》《数理统计与管理》等,文章质量普遍较高,作为《统计与精算》杂志的重要来源刊,长期稳定地为该刊供稿,是刊物质量的保证。其次,相关学科的期刊如《数量经济技术经济研究》《保险研究》,也为该刊提供了大量优秀

的文章,是刊物内容的重要支撑。再次,一些非统计类期刊如《经济研究》《农业技术经济》等,也刊登了大量运用统计方法进行研究分析的文章,无论是数据的获取以及处理,还是统计模型的采用,都体现出较高的水平,因此被该刊广泛选用。

(二)知识图谱

1. 研究热点

基于2014—2019年转载文章正文的文本,使用统计与精算学科词汇词典分词并去除停用词后,计算各词语在每篇文章中出现的词频(TF)与所有文章中包含该词语的逆向文件频率(IDF),用两者的乘积TF-IDF作为各文章中的词语的权重。在剔除单个词语后,计算各有效词语在所有文章中的TF-IDF权重之和,以此作为各词的重要程度的度量,避免了直接使用词频导致的对某些常用但意义不大的词语的重要程度的过高估计。在此结果上对统计与精算学科研究领域的整体热点与分类热点进行可视化。

(1)整体研究热点

利用TF-IDF对统计与精算学科的整体研究热点进行关键词抽取,从图3的前100个研究热点词云中可看到,研究热点集中在排放、家庭、农民工、核算、风险、CPI、贫困、收入、企业、教育等话题。



图3 统计与精算学科2014—2019年整体研究热点图

(2)分类研究热点

《统计与精算》杂志在常设栏目之外,每年还会策划1—2个专题,聚焦经济社会和学科发展中的重要问题,如“大数据时代下的统计学”“全要素生产率研究”“农村社会变迁”“CPI研究”等。为对《统计与精算》转载的有关统计类和精算类的文章作分类别的研究热点分析,本文按栏目和专题把转

载的文章进行归类,且在研究分类的热点时选择出现文章数大于4篇的词汇。

结合刊物定位,将CPI研究、全要素生产率研究、公共服务均等化、农村社会变迁、经济增长方式转型、产业活动统计、人口与人力资源统计、住户活动统计、公共部门统计、农业统计、国民经济核算、宏观经济分析、对外经济统计、收入分配统计、环境与生态统计、社会发展统计、医疗卫生统计、空间统计、比较与借鉴、学科建设、综论、产业活动统计(工业统计)划分为社会经济统计类别,合计文章251篇,占比67.47%;大数据时代下统计学、大数据时代下的统计学、理论与方法、研究动态、金融统计、金融统计(数理金融)等划分为数理统计类别,合计

文章47篇,占比12.63%;保险精算、风险管理划分为风险管理与精算类别,合计文章74篇,占比19.90%。社会经济统计类、数理统计类和精算类的研究热点(见图4)。

社会经济统计的研究热点集中在排放、家庭、农民工、核算、贫困、教育、收入、生产率、贸易、出口、增加值、企业、CPI、要素、支出上;数理统计的研究热点集中在关键词、搜索、估计量、统计学、季节、CPI、预测、网络、算法、抽样框、互联网、空间、惩罚、贝叶斯、异常值上;风险管理与精算的研究热点集中在风险、Copula、死亡率、索赔、长寿、灾害、预测、保险、损失、违约、跳跃、利率、银行、VaR、巨灾,各热点词的词频、出现文章数、TF-IDF信息(见表2)。

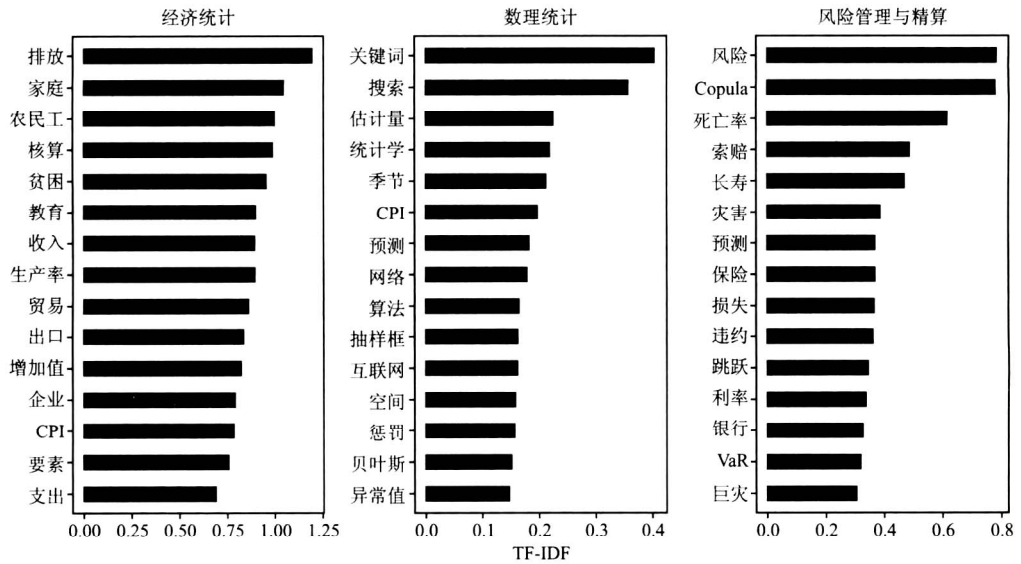


图4 社会经济统计、数理统计、风险管理与精算学科 Top15 研究热点图

表2 社会经济统计、数理统计、风险管理与精算学科 Top15 研究热点词信息

序号	社会经济统计				数理统计				风险管理与精算			
	词语	频数	文章数	TF-IDF	词语	频数	文章数	TF-IDF	词语	频数	文章数	TF-IDF
1	排放	1 317	27	1.20	关键词	276	9	0.40	风险	2 553	67	0.79
2	家庭	3 442	111	1.05	搜索	380	15	0.36	Copula	393	10	0.78
3	农民工	1 221	41	1.01	估计量	370	16	0.23	死亡率	728	15	0.62
4	核算	2 410	96	0.99	统计学	219	17	0.22	索赔	351	11	0.49
5	贫困	1 245	33	0.96	CPI	260	10	0.20	长寿	470	9	0.47
6	教育	3 286	135	0.90	预测	496	32	0.18	灾害	251	8	0.39
7	收入	4 925	176	0.90	网络	357	24	0.18	预测	902	51	0.37
8	生产率	1 940	72	0.90	算法	247	26	0.17	保险	471	39	0.37

续表 2

序号	社会经济统计				数理统计				风险管理与精算			
	词语	频数	文章数	TF-IDF	词语	频数	文章数	TF-IDF	词语	频数	文章数	TF-IDF
9	贸易	1 459	68	0.87	互联网	232	20	0.16	损失	714	49	0.37
10	出口	1 427	55	0.84	空间	660	21	0.16	违约	363	13	0.37
11	增加值	1 614	73	0.83	惩罚	144	12	0.16	跳跃	195	6	0.35
12	企业	3 833	139	0.80	贝叶斯	152	12	0.15	利率	593	27	0.34
13	CPI	999	27	0.79	异常值	67	10	0.15	银行	461	22	0.33
14	要素	2 577	122	0.76	高频	147	9	0.14	VaR	266	21	0.32
15	支出	2 532	122	0.70	属性	127	13	0.13	巨灾	223	8	0.31

2. 学科知识图谱及演化

本部分利用自然语言处理的方式从文章的标题、关键词、摘要中提取所有文章的名词性术语，并利用词语的共同出现情况生成《统计与精算》近 6 年转载文章的主题共现网络(见图 5)。由于用于分析的文献数量较少，如果按照年份划分成不同时间切片会导致结果较为稀疏，所以只分成一个时间切片进行整体的分析。在对主题词进行聚类并从关键词中提取聚类名称作为研究热点后发现，近年来统计与精算领域的研究热点主要集中在：

地方政府竞争: 主要涵盖投资竞争、城乡收入差距、人力资本、假设检验、劳动收入占比、资本深化、财政分权、农村劳动力迁移等方向，该类的文献较早年份为 2013 年。

全要素生产率: 主要和规模效应、生产前沿法、投入产出分析、增长核算、引力模型、能源消耗、空间杜宾模型、产业结构、结构效应、gvar 模型等共同出现，该类文献较早年份为 2014 年。

经济增长: 主要和 ces 生产函数、要素分配参数、技术进步、基尼系数、问卷调查、收入差距、经济禀赋、官员特征、收入分配、公共支出、居民幸福感、偿付能力、技术进步等共同出现，该类文献较早年份为 2014 年。

住户调查: 主要和固定资产投资、统计改革、金融生态、系统动态、公共服务、收入差距、资金流量账户、资金流量分析、统计外收入、保险消费等共同出现，该类文献较早年份为 2015 年。

大数据: 主要和互联网大数据、政府统计、应用



图 5 主题共现分析图

路径、动态因子模型、大数据计算机、数据质量、收缩估计、统计学理论、机器学习、创新应用、门限回归、动态因子模型、CPI 舆情指数、混频数据回归等共同出现,该类文献的较早年份为 2015 年。

DEA:主要和国际贸易、技术溢出效应、非竞争型投入产出模型、时空演变、农业全要素生产率、中国县域、生产效率、技术发展、渠道分析、装备制造等共同出现,该类文献的较早年份为 2014 年。

分位数回归:主要和驱动因素、LMDI、结构变动、位置坐标、单位名录库、地理信息、AR 模型、动态 VAR、渐近正态、WCQR 估计、驱动因素等同时出现,该类文献的较早年份为 2014 年。

其他的研究方向如长寿风险、贸易增加值等因共现词的出现频率较低故不在此展开。

从图 6 的主题时间线可以看到,随着时间的演变发展,2013 - 2014 年学者主要在投入产出、收入、经济增长、能源消耗等如何对经济发展进行评价等统计问题及偿二代等精算问题中展开多层次的研究,而 2015 - 2016 年随着大数据、互联网及多种类型数据的普及,数理统计的方法如分位数回归、半参数估计、空间面板模型等数理统计问题也开始焕发活力,与此同时,经济问题开始集中在全要素生产率、技术进步等问题的研究上,并开始拓展到如何与新时代的变化进行结合的统计分析,2017 - 2018 年研究热点主要集中在数据质量、时空演变、收入分布等问题。

3. 科研合作网络

为发掘各机构的学者及研究者之间的高效科研合作模式,本部分从机构与个人两个层面深入研究统计与精算领域的网络合作关系,并对其中有较高影响力的机构或学者进行展示,从而反映学科当前发展的科研合作特点。

对各转载论文作者的机构进行合作网络构建(见图 7)后,发现 445 个研究机构中,64 个机构为独立研究性机构,其他的 381 个有合作的机构中的平均合作机构数为 1.63,且转载篇数较多的机构均选择与其他机构合作的方式进行创新,如中国人民大学统计学院、中国人民大学应用统计研究中心、厦门大学经济学院、山西财经大学统计学院、湖南大学金融与统计学院、武汉大学经济与管理学院、中国社会科学院数量经济与技术经济研究所转载量均超过 6 篇且其合作机构数均超过 5 家。

对研究机构的合作情况进行社会网络分析可发现,中国人民大学统计学院、中国人民大学应用统计研究中心、厦门大学经济学院统计系、厦门大学经济学院、上海财经大学统计与管理学院、北京大学经济学院、西南财经大学统计学院、福建省统计科学重点实验室(厦门大学)的合作机构数均超过 7 家,国家统计局统计科学研究所、厦门大学经济学院统计系、东北财经大学统计学院、厦门大学经济学院、中国人民大学统计学院、厦门大学王亚南经济研究院、北京大学光华管理学院、华南师范大

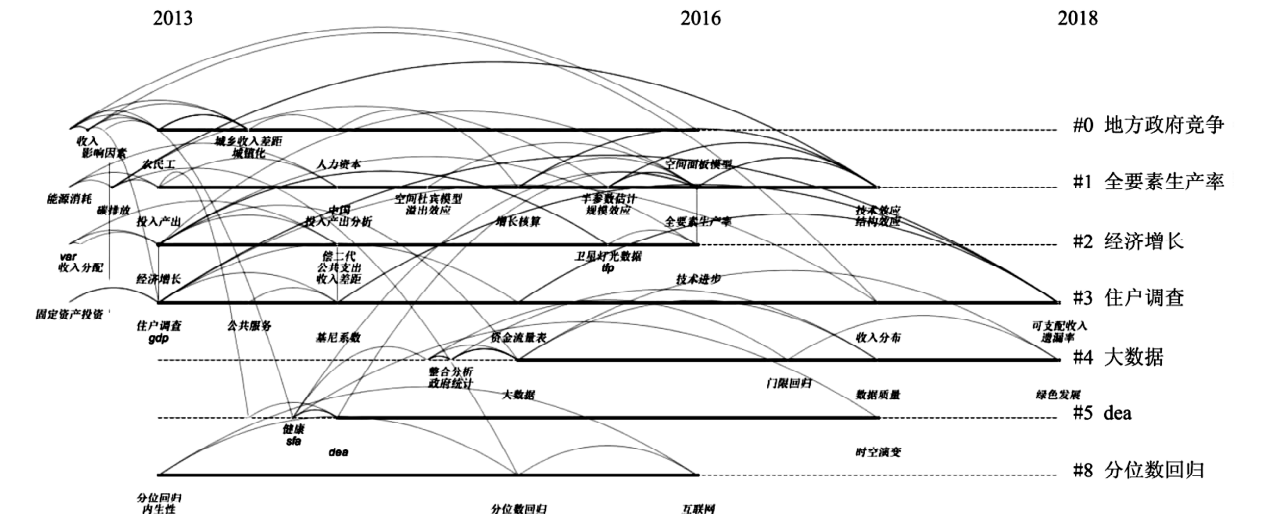


图 6 主题时间线图



图7 重要机构合作网络图

学经济与管理学院的中心度较高表明其在合作网络中发挥着中流砥柱的作用,中国人民大学统计学院、中国人民大学应用统计研究中心、厦门大学经济学院、厦门大学经济学院统计系、上海财经大学统计与管理学院、北京大学经济学院、西南财经大学统计学院、湖南大学金融与统计学院、武汉大学经济与管理学院、清华大学经济管理学院均具有较高的 Pagerank 值,表明其在统计与精算领域的研究中具有较重要的整体地位。

全部转载的 372 篇文章中,仅 60 篇为独著,特别是风险管理与精算类型中独著文章比例仅有 8.11%,社会经济统计、数理统计的独著比为 17.93%、18.75%。风险管理与精算中 52.70% 的文章为 2 人合作,超过半数。社会经济统计中 41.83% 的文章为 2 人合作,33.86% 的文章为 3 人合作;数理统计中 41.67% 的文章为 3 人合作,

35.42% 的文章为 2 人合作,也出现了 1 篇 6 人合作模式的高质量论文。这表明 2 至 3 人的合作模式为统计与精算文章较常采用的合作模式(见图 8)。

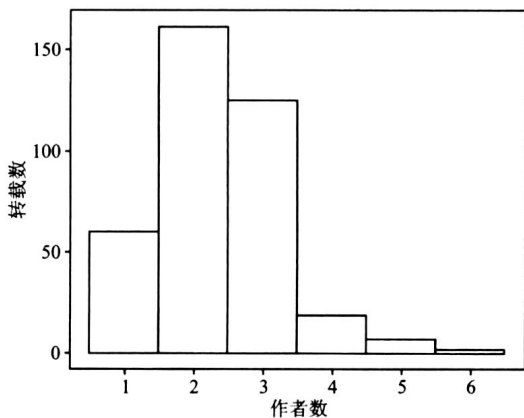


图8 论文作者数频数直方图

结合 6 年中转载篇数在 3 篇以上作者的统计数据可知,孟生旺、许宪春、金勇进、赵彦云、高敏雪、

表 3 重要研究机构合作网络信息

机构	频次	度数	中心度	PageRank	最早年份
中国人民大学统计学院	29	31	0.05	7.66	2013
中国人民大学应用统计研究中心	21	22	0.03	5.33	2013
厦门大学经济学院	11	12	0.08	3.46	2014
山西财经大学统计学院	9	6	0.01	1.74	2014
湖南大学金融与统计学院	8	7	0.01	2.28	2014
武汉大学经济与管理学院	8	6	0.03	2.19	2014
中国社会科学院数量经济与技术经济研究所	7	5	0	2.04	2015
国家统计局统计科学研究所	6	7	0.1	2.1	2013
清华大学经济管理学院	6	7	0.02	2.19	2014
北京大学经济学院	6	8	0.04	2.58	2014
上海财经大学统计与管理学院	5	9	0.02	2.84	2015
国家统计局	5	2	0	1	2013
厦门大学管理学院	5	3	0	1.03	2015
西南财经大学统计学院	5	8	0.01	2.57	2016
复旦大学经济学院	5	7	0.04	2.15	2014
厦门大学经济学院统计系	5	13	0.09	2.89	2014
对外经济贸易大学保险学院	5	6	0.02	1.62	2014
东北财经大学统计学院	4	5	0.09	1.37	2014
中山大学岭南学院	4	3	0	1.3	2015
北京师范大学经济与工商管理学院	4	5	0	2.06	2013
上海交通大学安泰经济与管理学院	4	2	0	1.06	2014
北京航空航天大学经济管理学院	4	4	0.01	1.58	2015
浙江工商大学统计与数学学院	4	3	0	1.79	2016
北京师范大学统计学院	4	1	0	1	2015
吉林大学数量经济研究中心	4	3	0	1.92	2013
中国人民大学农业与农村发展学院	4	4	0	1.89	2016
兰州财经大学统计学院	4	5	0.06	1.32	2016

注:重要研究机构指频次 ≥ 3 、度数 ≥ 1 的机构。

张连增等 16 位作者 6 年共被转载 57 篇文章(见表 4),他们是该刊的重要作者,这些学科带头人在作者合作网络中较为分散,分别带领所在团队在不同领域创造出高质量的研究成果,48 篇合作成果中孟生旺、金勇进、高敏雪、张连增的合作成果数均超过

3 篇。朱建平、马双鸽在网络中的 Pagerank 值最高,表明在合作网络中具有重要作用,度数较高的除此二位学者外还包括宫晓琳、张宁、杨淑振。学科带头人的引领作用与高能力研究者的自发合作共同促进着统计与精算学科的繁荣发展(见第 121 页图 9)。

表 4 2014 - 2019 年转载 3 篇及以上论文的作者转载信息

序号	作者姓名	第一作者单位	文章题目	原发期刊	转载期刊
1	孟生旺,王明高	中国人民大学统计学院	非寿险未决赔款准备金评估的增长曲线模型	经济管理	2015.1
2	刘新红,孟生旺	中国人民大学统计学院	相依风险条件下的医疗费用预测模型及其应用	数理统计与管理	2016.1
3	王明高,孟生旺	中国人民大学统计学院	尖峰厚尾巨灾损失数据的组合分布模型	保险研究	2018.2
4	高光远,孟生旺	中国人民大学统计学院	基于车联网大数据的车险费率因子分析	保险研究	2018.3
5	李政宵,孟生旺	中国人民大学统计学院	相依风险的团体健康保险损失预测	数理统计与管理	2018.5
6	许宪春	国家统计局国际经济核算司	中国当前重点统计领域的改革	经济研究	2014.1
7	许宪春	国家统计局国际经济核算司	论中国国民经济核算体系 2015 年的修订	中国社会科学	2016.3
8	许宪春,郑学工	国家统计局国际经济核算司	改革研发支出核算方法更好地反映创新驱动作用	国家行政学院学报	2017.2
9	陈丹丹,许宪春,汤美微,余航	西南财经大学	基于保密视角的政府微观调查数据开发应用研究	经济纵横	2018.4
10	许宪春,任雪,常子豪	清华大学经济管理学院	大数据与绿色发展	中国工业经济	2019.4
11	戴明锋,金勇进,孙婕	中国人民大学统计学院	我国公民权利意识的抽样调查方案设计	数理统计与管理	2015.6
12	李毅,刘雅楠,金勇进	中国人民大学统计学院	基于复杂网络视角下省际人口迁移空间格局及趋势研究	统计研究	2018.1
13	万舒晨,金勇进	中国人民大学统计学院	大数据背景下小微企业多重抽样框应用研究	数量经济技术经济研究	2018.6
14	于力超,金勇进	中央民族大学	基于分层模型的缺失数据插补方法研究	统计研究	2019.2
15	金勇进,姜天英	中国人民大学	工业企业规模标准划分问题研究	统计与信息论坛	2019.6
16	刘思明,侯鹏,赵彦云	中国人民大学统计学院	知识产权保护与中国工业创新能力:来自省级大中型工业企业面板数据的实证研究	数量经济技术经济研究	2015.4
17	赵彦云	中国人民大学统计学院	对大数据统计设计的思考	统计研究	2015.5
18	曾相崑,赵彦云,贺飞燕	中国人民大学统计学院	中国人口的多边流动与再分布均衡	调研世界	2016.2
19	赵彦云	中国人民大学统计学院	加速开拓统计学的应用价值:中国统计教育回顾与展望	统计与信息论坛	2016.5
20	高敏雪,葛金梅	中国人民大学统计学院	出口贸易增加值测算的微观基础	统计研究	2014.1
21	高敏雪	中国人民大学统计学院	《环境经济核算体系(2012)》发布对实施环境经济核算的意义	中国人民大学学报	2016.2
22	高敏雪,李璐	中国人民大学统计学院	企业填报负担与政府统计指标开发利用程度:关于政府统计效率和公开性的案例研究	调研世界	2016.3
23	高敏雪,刘茜,黎煜坤	中国人民大学统计学院	在 SNA - SEEA - SEEA/EEA 链条上认识生态系统核算:《实验性生态系统核算》文本解析与延伸讨论	统计研究	2018.6
24	张连增,曲珩	南开大学经济学院	欧盟与美国 ORSA 研究及对我国的实践指导	保险研究	2014.1
25	孙维伟,张连增	南开大学金融学院	车险费率厘定精算技术的研究与应用评述	数理统计与管理	2017.4
26	张连增,王缔	南开大学金融学院	保险大数据条件下车险费率厘定的研究:基于 SOM 神经网络方法的车险索赔强度建模	保险研究	2019.1
27	张连增,申晴	南开大学金融学院	提升算法对传统车险索赔频率建模模型的改进:基于我国五省交强险保单数据	保险研究	2019.6
28	彭小辉,王常伟,史清华	南京师范大学	城市农民工生命统计价值研究:基于改进的特征工资模型——来自上海的证据	经济理论与经济管理	2014.3

续表 4

序号	作者姓名	第一作者单位	文章题目	原发期刊	转载刊期
29	程名望,史清华,许洁	同济大学	流动性转移与永久性迁移:影响因素及比较—— 基于上海市 1446 份农民工样本的实证分析	外国经济与管理	2014.5
30	刘彬彬,崔菲菲,史清华	上海交通大学	劳动力流动与村庄离婚率	中国农村经济	2019.1
31	马双鸽,王小燕,方匡南	美国耶鲁大学	大数据的整合分析方法	统计研究	2016.2
32	李扬,许文甫,马双鸽	中国人民大学	污染数据的稳健稀疏成组变量选择方法研究	统计与信息论坛	2018.5
33	“大数据中的统计方法” 课题组,马双鸽	厦门大学	大数据时代统计学发展的若干问题	统计研究	2017.3
34	魏瑾瑞,朱建平,谢邦昌	东北财经大学	金融高频数据仅仅是一个优质的时间序列吗: 概念及统计特征的再考察	投资研究	2014.6
35	许永洪,朱建平	厦门大学	消费者信心指数的偏差和修正研究	数理统计与管理	2018.6
36	许永洪,萧珍丽,朱建平	厦门大学经济学院统计系	教育缓解了收入分配不平衡吗	数理统计与管理	2019.5
37	许永洪,洪昕,陈建伟	厦门大学经济学院	与能源消耗等指标发生偏离的 GDP 是否可信	统计研究	2017.5
38	许永洪,朱建平	厦门大学经济学院	消费者信心指数的偏差和修正研究	数理统计与管理	2018.6
39	许永洪,萧珍丽,朱建平	厦门大学经济学院	教育缓解了收入分配不平衡吗	数理统计与管理	2019.5
40	李锐,傅小燕,向书坚	中南财经政法大学	养老金制度变革的福利损益与再分配研究	统计研究	2014.6
41	向书坚,温婷	中南财经政法大学统计与 数学学院	中国对外贸易隐含碳排放的重估算:基于新附加 值贸易统计视角	国际经贸探索	2015.2
42	向书坚,吴文君	中南财经政法大学统计与 数学学院	OECD 数字经济核算研究最新动态及其启示	统计研究	2019.3
43	王霞,洪永森	中国科学院大学管理学院	一类基于非参数回归的条件异方差检验	统计研究	2015.2
44	洪永森	发展中国家科学院	经济统计学与计量经济学等相关学科的关系及 发展前景	统计研究	2016.5
45	张兴祥,钟威,洪永森	发展中国家科学院	国民幸福感的指标体系构建与影响因素分析: 基于 LASSO 的筛选方法	统计研究	2019.2
46	赵明,王晓军	中国人民大学	保险公司长寿风险度量	统计研究	2016.2
47	王志刚,王晓军,张学斌	内蒙古财经大学	封闭人群生存人数整体预测——以 Lee - Carter 模型为例	人口与经济	2016.6
48	米海杰,王晓军,王述珍	中国劳动和社会保障科学 研究院	养老保险个人账户参数设定错误与纠正	统计研究	2019.5
49	段白鸽	复旦大学经济学院	动态死亡率建模与长寿风险量化研究评述	保险研究	2015.4
50	段白鸽,陆婧文	复旦大学经济学院	长寿风险对保险公司年金产品定价的影响:基于 区块 Bootstrap 方法的实证分析	山西财经大学学报	2015.6
51	段白鸽	复旦大学经济学院	长寿风险对寿险和年金产品定价的对冲效应研究	保险研究	2019.6
52	吕光明,于学霆,徐曼	北京师范大学国民核算 研究院	多方法改进视角下的中国核心通货膨胀测算和效 果评价	统计研究	2014.4
53	吕光明	北京师范大学统计学院	我国劳动报酬占比测算中的几个数据质量问题	商业经济与管理	2015.5

续表 4

序号	作者姓名	第一作者单位	文章题目	原发期刊	转载期刊
54	吕光明,于学霆	北京师范大学国民核算研 究院	我国食品价格上涨对总体价格影响的“二轮效 应”分析	经济统计学	2017.2
55	谢远涛,杨娟	对外经济贸易大学保险学院	基于操作时间和广义线性混合模型的准备金评估 技术研究	保险研究	2014.4
56	谢远涛,杨娟,刘皓宇	对外经济贸易大学保险学院	基于改进的自适应传播模型的农业风险区划分析	统计与信息论坛	2017.3
57	谢远涛,毛羽	对外经济贸易大学保险学院	基于 Gamma 随机效应分层贝叶斯模型的汽车延 保定价	统计与信息论坛	2018.3

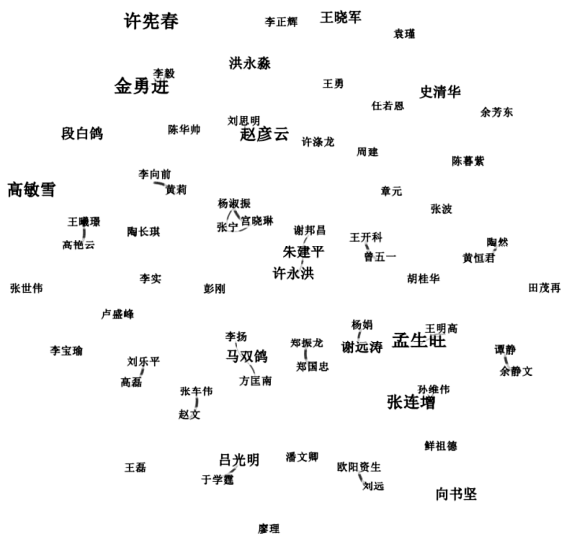


图9 重要作者合作网络图

四、结论

本文基于复印报刊资料《统计与精算》2014 - 2019 年转载的文章进行实证研究,通过文章类型、文章长度、文章来源深入分析学科发文特点,采用自然语言处理和文献计量的分析方法,探究学科热点、学科现状及演化、科研合作中的机构和作者的合作网络特点,并进行知识图谱的结构展现。结果显示,《统计与精算》杂志长期关注中国经济社会发展中的重大现实问题,以问题为导向,侧重研究问题的时代性和前沿性;统计与精算学科的发文呈现理论与实证相结合、趋向中篇幅聚拢的特点;专业期刊通过集聚效应在推动统计和精算学科的发展和学术成果的传播上发挥了重要作用,是学术研究的主阵地,非专业期刊特别是高质量的学术期刊,通过自身平台的影响力,组织编辑了大量优质稿件,是学科发展的有益补充;统计与精算学科整体聚焦在能源环境、特殊群体、消费、收入、教育等

民生问题及企业发展等社会经济问题,数理统计主要针对新数据特点(如高频、网络、互联网)、新问题(如搜索、预测、异常值)下的算法(如贝叶斯、空间分析)及统计量的理论性质进行创新,风险管理与精算重点关注寿险、非寿险的风险预测;随着时间的演变,统计与精算的研究热点从经济发展评价、偿二代到大数据、互联网下数理统计方法的创新,再到生产效率、技术进步,以及结合新时代的变化进行统计分析,近 2 年主要聚焦在数据质量、时空演变、收入分布等问题;高校和研究机构是统计与精算学科发展研究的重镇,特别是高等院校,占有绝对的比例,且发文数量和学校统计学科的质量相一致;学科带头人、专家团队在不同领域中带领所在团队进行学术研究,高能力研究者通过自发合作为学科发展带来新的活力,共同引领着统计与精算学科的发展;研究机构采取 2—3 名学者合作发文的模式特征明显,在相关学术研究越来越专业化的趋势下,通过合作的形式达成优势互补、资源共享,有利于高质量成果的产出。

基于以上研究结论,统计与精算学科领域学者应紧密结合当前互联网时代下的新数据、新场景进行创新,致力于为社会发展中的重大问题建言献策。同时,随着高校向研究型大学转型,不同类型的高校在科研实力上的差距正逐步拉大,不同机构之间学者的合作必将成为实现双赢的重要途径,学科带头人在引领学科发展的同时应为培育学科发展的新兴力量做出自己的贡献^[4-7]。

研究结果主要依据复印报刊资料《统计与精算》近 6 年的全文转载数据得出,周期较短、数据量较小,仅代表应用统计和精算学科近年来高质

量研究中反映出的问题,或可起到为相关研究人员和对复印报刊资料长期支持、关注的朋友们提供参考的作用。在统计学的发展过程中,始终对经济社会发挥着重要的作用,并表现出强大的生命力^[8]。面对大数据时代的来临和新一轮信息技术革命的挑战,统计学人只有适应变化并不断探索、勇于创新,方能历久弥新,在构建新时代中国特色哲学社会科学话语体系中担负起自己的责任和使命。

参考文献:

[1] 复印报刊资料重要转载来源作者(2019年版)研制报告[R]. 中国人民大学学术成果评价研究中心,2019.

[2] 复印报刊资料重要转载来源机构(2018年版)研制报告[R]. 中国人民大学学术成果评价研究中心,2018.

[3] 复印报刊资料重要转载来源期刊(2017年版)研制报告[R]. 中国人民大学学术成果评价研究中心,2017.

[4] 赵彦云. 例践透视中国统计40年[J]. 统计与信息论坛,2019,34(5):118-128.

[5] 杜文胜. 中国经济统计学70年:学科发展发展历程与前景展望[J]. 经济研究参考,2019(20):49-63.

[6] 游士兵,黄靖义. 量化史学视角下中国统计组织史研究[J]. 统计与信息论坛,2019,34(12):117-124.

[7] 郜泽玲,邢春冰. 经济学中文权威期刊的几个典型事实——以《经济研究》和《管理世界》为例[J]. 经济资料译丛,2019(3):59-71.

[8] 赵彦云,车明佳. 统计能肩负互联网时代发展的伟大使命吗?[J]. 中国软科学,2020(5):56-62.

Bibliometrics Analysis of the Reprinted Articles of *Statistics and Actuarial Science* from 2014 to 2019

Zhang Hao Che Mingjia

Abstract: Taking the reprinted articles of *Statistics and Actuarial Science* from 2014 to 2019 as the research object, through the word frequency analysis of the reprinted high-quality papers, the overall research hot-spots of statistics and actuarial science in the past six years are explored by using bibliometrics software Citespace 5.7 R1 to map the co-word information of the title, abstract and key words of papers, finds the transformation of knowledge structure and research trends, and reveals the law of academic cooperation in this field based on the research cooperation network of the authors of *Statistics and Actuarial Science*. The results show as follows: the journal is problem-oriented, focusing on the epochal and cutting-edge of research issues; universities and research institutions are important sources for the development of statistics and actuarial science; academic leaders and expert teams play a leading role in the discipline construction of their institutions; professional journals are the main front for the dissemination of academic achievements of statistics and precision computing; institutions with more papers choose to cooperate with other institutions; multiple author cooperation has become a feature of academic achievements published in the field of statistics and actuarial science, scholars achieve win-win results through complementary advantages and resource sharing. *Statistics and Actuarial Science* have made their own unique contributions to the promotion of discipline development and academic communication through extensive observation and selection of the best.

Key words: information center for social sciences of renmin university; industry; *Statistics and Actuarial Science*; bibliometrics; knowledge graph