



教育部人文社会科学重点研究基地重大项目成果丛书

Publication Series: MOE Supported Projects of Key Research Institutes of Humanities and Social Sciences in Universities

新闻传播学图书馆情报学档案学类

Journalism and Communication, Library
and Information Science, Archival Science

文献计量内容分析法

邱均平 王曰芬 等 著



国家图书馆出版社

教育部人文社会科学重点研究基地重大项目成果丛书

文献计量内容分析法

邱均平 王曰芬 等著

文献计量内容分析法 著者 邱均平, 王曰芬
出版 国家图书馆出版社
发行 010-66139742
E-mail cps@nlc.gov.cn

文献计量内容分析法 著者 邱均平, 王曰芬
出版 国家图书馆出版社
发行 010-66139742
E-mail cps@nlc.gov.cn

文献计量内容分析法 著者 邱均平, 王曰芬
出版 国家图书馆出版社
发行 010-66139742
E-mail cps@nlc.gov.cn

Website www.nlcpress.com

总发行 新华书店

印刷 北京汉玉印刷有限公司

开本 787 × 1092 毫米 1/16

印张 22.75

版次 2008 年 10 月第 1 版 2008 年 10 月第 1 次印刷

定价 600 (千字)

国家图书馆出版社

书号

定价 72.00 元

图书在版编目(CIP)数据

文献计量内容分析法/邱均平,王曰芬等著. —北京:国家图书馆出版社,2008. 10

ISBN 978 - 7 - 5013 - 3547 - 3

I. 文… II. ①邱…②王… III. 文献计量学—内容分析 IV. G256

中国版本图书馆 CIP 数据核字(2007)第 182795 号

书名 文献计量内容分析法

著者 邱均平,王曰芬 等著

出版 国家图书馆出版社 (100034 北京西城区文津街7号)

发行 010 - 66139745 66151313 66175620 66126153

66174391(传真) 66126156(门市部)

E-mail cbs@nlc.gov.cn(投稿) btsfxb@nlc.gov.cn(邮购)

Website www.nlcpress.com

经销 新华书店

印刷 北京汉玉印刷有限公司

开本 787 × 1092 毫米 1/16

印张 25.75

版次 2008 年 10 月第 1 版 2008 年 10 月第 1 次印刷

字数 600(千字)

书号 ISBN 978 - 7 - 5013 - 3547 - 3/G · 731

定价 75.00 元

内容提要

文献计量分析法,是以文献信息为研究对象、以文献计量学为理论基础的一种研究方法。内容分析法,是对文献内容进行客观、系统和量化描述与分析的研究方法,是社会科学研究中普遍使用的一种科学方法。本书力图对二者进行全面比较分析,将这两种分析方法结合起来,构建一种新的分析研究方法——文献计量内容分析法(简称“综合方法”)。全书共分八章,内容包括理论基础研究、软件工具研究、比较研究、综合研究、技术实现研究、应用研究等,重点探讨了综合方法的方法论依据和理论基础、方法体系构建,以及实际应用案例与实证分析等。最后对综合方法应用中的主要影响因素进行了讨论。本书受教育部人文社会科学重点研究基地基金资助,作为一本关于科学研究方法的专著成果,有较强的创新性、前瞻性与可操作性,既可为文献计量与内容分析法的研究提供有价值的参考和启示,也可作为科学研究工作者的方法指导与参考用书。

前言

经过几十年的研究和发展,文献计量学(Bibliometrics)的理论和被揭示的六大规律已经成为图书情报学的重要基础和组成部分,同时也形成了一些较为成熟的方法,如文献统计分析法、书目分析法、引文分析法等。这些分析方法目前都是文献计量学直接应用的重要方法,它们共同构成了文献计量分析法,已发展成为图书情报学不可缺少的专门的特征方法;其应用研究也开始由图书情报领域拓展到科学学、社会学、预测学、科学评价和科技管理等许多领域,为科学管理和决策提供定量依据,并发挥重要作用,因而越来越受到国际学术界、科学界和政府管理部门的高度重视和大力支持。网络信息计量学(Webometrics)正在兴起,已迅速成为人们普遍关注和研究的热点,电子化、数字化、网络化环境下的文献计量分析正在成为新的研究领域;文献计量与内容分析等方法的交叉、结合和综合运用更是国际学术界普遍重视的前沿性重大课题,亟待研究和解决。我国对文献计量学的研究虽然起步较晚,但发展很快,目前已经初具规模,基本上形成了研究、教育和实际应用全面发展的良好局面,并不断取得新的进展,已成为图书情报与科学评价领域中一个重要的分支学科,不仅得到了国内外图书情报界的普遍承认和重视,而且还逐步被社会所接受,在整个社会科学研究和科学管理及决策中的作用也日益显现出来。但是,与国外发展水平相比,仍有一定差距,其表现之一就是:如何综合利用文献计量分析法与内容分析法,深入到文献内部对其内容、知识进行测度和综合研究,这是图书情报学和信息管理领域当前面临的重大课题之一。

从内容分析法(Content Analysis)是一种对文献内容进行客观、系统和量化描述与分析的研究方法,是社会科学研究中普遍使用的一种科学方法。对文献内容的研究可分为定性与定量两种方式,分别从不同侧面对文献中所包含的信息和知识进行加工处理与分析,从而得出相应的研究结论。但人们通常将对文献内容的定量研究称为内容分析法。内容分析法最早萌发于新闻界,后来扩展到图书情报乃至整个社会科学领域,在信息传播与情报研究中有着重要的地位和作用。目前,国内外有关内容分析的研究主要以词频统计分析为主,存在一些局限。但近些年来,由于计算机技术应用于内容分析领域,对内容分析的发展产生了根本性影响;全

文数据库的发展又为内容分析提供了便利条件;数据挖掘和知识发现技术的应用也给内容分析方法研究带来了蓬勃生机。当前,如何充分利用数据库系统和网络信息资源,实现内容分析的自动化,提高研究效率和分析结果的可靠性与准确性是目前急需重点研究和解决的重大课题。

事实上,以上两种方法都有各自的优点与局限,而目前的文献情报研究又大多是采用单一方法进行的,因而影响了研究结论的准确性和可信度。为了克服采用单一方法研究造成的弊端和局限,必须将文献计量分析法与内容分析法结合起来开展综合研究,以便提高文献信息与内容研究的效率和分析结果的准确性。这是当前国际学科前沿研究的重大课题和新的趋势。但目前已发表的有关论文很少,已有的也只是一般的方法介绍,或者是两种方法的简单比较,至今还没有全面、系统的深入研究。因此,文献计量分析法与内容分析法的“综合研究”仍是一个全新的学科前沿领域,是一个新的学科生长点和研究热点,我们应该尽快加强这方面的研究。因此,本课题的研究具有重要的科学理论价值和现实指导意义。

本书共分为八章,从方法论角度对文献计量分析法与内容分析法进行了综合研究,其内容主要包括以下几个方面:理论基础研究、比较研究、综合方法研究、应用软件与工具的开发研究、综合应用研究。我们力图在对这两种方法进行详细比较研究的基础上,将文献计量分析法与内容分析法结合起来,取长补短,构建一种新的分析研究方法——文献计量内容分析法(简称“综合方法”);重点研究其方法论依据和理论基础,方法的原理、基本流程与模式,以及技术实现与实际应用的范围、应用案例和主要影响因素等。在比较研究部分,主要是从总体上比较这两种分析方法的相似性与差异性,系统研究文献计量分析法与内容分析法在基本构成,方法论基础,理论原理,方法模式与基本程序,适用的对象、范围和条件,应用案例,发展趋势等方面的异同点,弄清两者各自的特点、优势与局限,为综合方法的论证奠定基础。综合研究部分是在比较研究的基础上,从方法论出发,研究文献计量分析法与内容分析法结合的必要性和可行性,找出和理清两者综合的切入点与思路;探索两者综合研究的方法论依据与基础,论证综合研究对象可提取的信息特征,揭示综合研究的计量分析单元隐含的知识价值,构建综合研究的方法途径和体系架构,研究综合方法应用的基本程序与应用的范围等。最后,分别选择基于时空分布维度和内容层次维度的综合方法的应用途径,结合实际应用案例进行综合方法的应用与评价研究,并分析其优缺点以及影响综合方法应用的主要因素等。

本书是教育部人文社会科学重点研究基地重大项目“文献计量与内容分析的综合研究”(项目号:02JAZJD870003)的最终研究成果。本课题由邱均平教授负责主持。近五年来,我们课题组全体成员在邱均平教授的带领下,从不同角度不同方面展开了全方位的系统研究,先后公开发表研究论文30余篇,有的被人大报刊复印资料全文转载或被引率较高。2004年11月,我们在武汉大学专门召开了“第三届信息化与信息资源管理学术研讨会”,对这个重大项目进行了专题研讨,会后还从来自全国的大量论文中精选34篇正式出版了《文献计量与内容分析》论文集,在学术界产生了广泛影响。邹非硕士、王曰芬博士还针对本课题撰写了学位论文,获得了评审专家的好评。本书就是在以上阶段性成果的基础上总结、补充和修改而成的,同时还参考了一些国内外其他的相关文献。它既展示了本课题研究的丰硕成果和水平,同时也为进一步推动网络环境下文献计量与内容分析法的理论基础和方法论研究以及实际应用提供了有价值的参考和启示。我们希望本课题的研究成果和本书的出版能激发更多研究者的兴趣和积极性,共同推动本领域研究进一步向广度和深度发展。

本书由邱均平、王曰芬主撰。任全娥协助完成了全书的统稿工作,为本书的顺利出版作出了较大贡献。本书的出版得到了武汉大学信息资源研究中心、信息管理学院和南京理工大学信息管理系以及国家图书馆出版社有关领导的支持和帮助,特别是宋安莉主任给予了大力支持,王涛为本书的编辑工作付出了大量的辛勤劳动。在此,我们一并表示由衷的感谢!

邱均平

于武汉大学 2007 年 7 月 9 日

(845)		3.0
(825)		5
(828)		1.7
(505)		5.7
(153)		8.7
1	文献计量分析法与内容分析法概论	(1)
1.1	文献计量分析法的概念和研究进展	(1)
1.2	内容分析法的概念和研究进展	(10)
2	文献计量分析法与内容分析法的理论基础	(25)
2.1	文献计量分析法的理论基础	(25)
2.2	内容分析法的基本理论问题探讨	(32)
2.3	内容分析法的理论构建	(43)
2.4	内容分析法中的不确定性	(54)
2.5	内容分析法与知识发现的比较分析	(62)
3	文献计量分析法与内容分析法的软件工具	(70)
3.1	文献计量分析法的软件工具分析	(70)
3.2	网络信息计量分析的软件工具	(82)
3.3	内容分析法的软件工具分析	(107)
3.4	计算机辅助内容分析软件工具的研究进展	(113)
4	文献计量分析法与内容分析法的比较研究	(120)
4.1	方法构成与特性的比较分析	(120)
4.2	方法论基础的比较分析	(125)
4.3	基本程序与实施机制的比较分析	(129)
4.4	实际应用的比较分析	(138)
4.5	最新研究与发展趋势的比较分析	(143)
4.6	两者相似性与差异性的综合比较分析	(153)
5	文献计量内容分析法的综合研究	(160)
5.1	综合研究的必要性、可行性和基本思路	(160)
5.2	综合研究的方法来源与方法构成的理论研究	(176)
5.3	综合研究的方法体系构建	(197)
6	文献计量内容分析法的技术实现	(219)
6.1	文献计量分析法的技术实现	(219)
6.2	内容分析法的技术实现与汉语自动分词	(227)
6.2	基于网络的内容分析法及其应用	(239)

6.3	文献计量内容分析法综合应用的技术实现	(248)
7	文献计量内容分析法的应用研究	(258)
7.1	文献计量分析法的应用研究	(258)
7.2	内容分析法的应用研究	(292)
7.3	文献计量内容分析法的综合应用	(321)
8.1	文献计量内容分析法发展中的影响因素	(381)
(1)	8.1 数据清洗	(382)
(10)	8.2 专业人员及其情报研究	(387)
(25)	8.3 信息可视化	(392)
	参考文献	(398)
(35)		3.5
(43)		3.5
(44)		4.5
(50)		2.5
(70)		3
(70)		1.3
(85)		2.3
(101)		3.3
(113)		3.4
(150)		4
(150)		1.4
(152)		2.4
(159)		3.4
(138)		4.4
(141)		4.2
(153)		4.4
(160)		2
(160)		1.2
(170)		2.2
(191)		3.2
(215)		6
(219)		1.6
(225)		2.6
(230)		2.6

1 文献计量分析法与内容分析法概论

1.1 文献计量分析法的概念和研究进展

1.1.1 文献计量分析法概述

我们把文献计量分析法界定为以文献信息为研究对象、以文献计量学为理论基础的一种研究方法。

在文献情报研究方法中,文献计量分析法是一种基于数学和统计学的著名的定量分析方法,它以各种科学文献的外部特征为研究对象,以输出量必定是量化的信息内容为主要特点,采用数学与统计方法来描述、评价和预测科学技术的现状与发展趋势的图书情报学分支学科。这些定量分析研究必须建立在对与所研究的目的有密切关系的媒介统计数字的基础上,只有在这种统计数字的基础上,才能分析和预测到学科文献的各种特征及其变化规律。文献计量分析法统计分析的对象可以是各类文献本身及其所表现出来的各种特征,如书名、著者、出版年、出版地、出版商、参考文献、页数、版次、价格及文献内容等。但事实上往往不止以上所述的几种,一切与文献有关的媒介及其特征都可以作为进行文献计量分析研究的对象,文献计量分析的统计对象是多方面的。

经过几十年的发展,文献计量分析法的理论和已揭示的六大定律已经成为情报学研究的重要基础和组成部分。在许多学者潜心研究文献规律性理论的同时,也形成了一些较为成熟的方法,如文献统计分析法、书目分析法、引文分析法等,其应用也由图书情报部门拓展到科学学、社会学、预测学、科学评价和科技管理等许多领域,为科学管理和决策提供定量依据,发挥着重要作用。

1.1.1.1 文献计量分析法的数据来源^①

文献计量分析法的数据来源随着时代的变化也有所变化,这里介绍的数据来源包括传统的“十大信息(情报)源”和近几年发展起来的光盘

^① 邱均平. 文献计量学. 北京:科学技术文献出版社,1988.

数据和网络数据等新型数据来源。

(1)科技图书:是对某专门知识或某学科进行系统的论述或概括的一种情报来源。科技图书包括专著、论文集、教科书、百科全书以及字典、词典和手册等。一般来说,科技图书的出版周期较长,对新成就、新观点、新技术等的介绍不如科技期刊来得快。

(2)科技期刊:又称连续性出版物(serials),是一种定期或不定期连续出版的文献载体,它一般有同样的名称,按一定时间编定卷号,每一卷又分若干期。科技期刊出版周期短,因而发表论文速度快;科技期刊种类多,因而能反映各个领域的成果和水平。查阅科技期刊,对于了解第一手资料、掌握进展情况、开阔思路最有参考价值。其类型有:A.学术性、技术性刊物,是科技期刊的核心部分。刊名一般为:Acta(学报)、Journals(杂志)、Annuals(纪事)、Bulletin(通报)、Transactions(汇刊)、Proceedings(会刊)、Review(评论)、Progress/Advances in...(进展)等。B.快报型。如Communication(通讯)、letters(通讯)、bulletin等。C.消息性(news journals)。news(新闻)、news letters等。D.资料性(data journals)。data、event等。

(3)科技报告:报道(记录)研究和开发调查工作的成果或进展情况的一种文献类型,一般都编有号码以供识别。常用的报告名称按报告的形式可分为:report(报告书)、technical notes(技术札记)、memorandum(备忘录)、papers(论文)、bulletin(通报)、technical translations(技术译丛)、special publications(特种出版物);按研究的进展情况分:primary report(初步报告)、progress report(进展报告)、interim report(期中(临时)报告)、final report(最终报告)。国际上比较著名的科技报告如美国的AD(国防部ASTIA Document)报告、NASA报告、PB报告。

(4)会议文献:是指学术会议文献,它往往反映出科学技术的发展趋势,其特点是与最新成果的间隔时间短,但其内容与期刊相比可能不太成熟。会议类型大致可分为国际会议、全国性会议及地区会议等。会议和会议文献常用的主要名称有conference(大会)、meeting(小型会议)、symposium(讨论会)、proceeding(会议录)、paper(单篇论文)、transaction(汇报)等。如美国冷泉港定量生物学讨论会(Cold Spring Harbor Symposia on Quantitative Biology)会议录、欧洲生化学会联合会(Federation of European Biochemical Society, FEBS)专题会议录等。

(5)专利文献:在实行专利制度的国家,凡是本国或外国的个人和企业有了创造发明,都可以根据专利法的规定,向本国或外国专利局提出申

请,经审查合格,批准授予在一定年限内享有创造发明成果的权利,并在法律上受到保护,这样一种受到法律保护的技术专有权利叫做专利。专利是与产业活动密切相关的,因而其实用性非常强。根据专利的技术水平和应用情况,其类型有 invention(发明)、utility patent(实用新型)、design patent(外观设计)等。《中华人民共和国专利法》自 1985 年 4 月 1 日起施行。

(6) 标准文献:标准化主要包括三个方面:①产品标准化,即产品质量要符合技术规定;②产品规格化和系列化,即产品按型号大小来分档,成系列的发展,从而保证以较少的品种来满足广泛的要求;③零部件通用化,即同类机型零部件,特别是易损件要做到最大限度的通用互换性。经过公认权威当局批准的标准化工作成果,可以采用文件形式或规定基本单位(物理常数)这两种形式固定下来,以文件形式出现的标准化工作成果,就是标准文献。

(7) 学位论文:是高等学校毕业生、研究生在评定学士、硕士、博士等学位时提出的论文。因为它是经过一定审查的原始研究成果,一般都带有独创性和学术性,所以也有一定的参考价值。学位论文一般不单独出版,质量比较高的论文,多发表在有关专业的期刊上。

(8) 产品资料:一般是指产品样本,即产品说明书。好的产品说明书含有丰富的内容,包括产品规格、产品特点、产品专利号等多种对生产有用的信息。

(9) 技术档案:是指在生产建设和科技部门的技术活动中形成的,有一定的工程对象的技术文件的总称。

(10) 科技报纸:在报纸上刊载的科技文献,主要是指关于科技方面的新闻报道。由于报纸报道及时、包含的信息量大,所以,它可使我们及时得到一些重要的科技消息。同时,报纸上的有关科技专栏,也是我们了解国内外科技动态和科技珍闻的一个窗口。

(11) 光盘数据:光盘存储的数据量大,数据组织有序,现在不少引文索引数据库都有光盘版,比如国内《中国科学引文索引》(CSCI)、《中国科技论文与引文数据库》(CSTPC)和《中文社会科学引文索引》(CSSCI),另外国外的 SCI、EI 等大型数据库都有光盘数据库。所有这些光盘数据对于我们进行信息检索和科研评价等都有重要的意义。

(12) 网络数据:随着网络的发展,越来越多的数据存放在互联网上,这些数据更新及时,查询非常方便,诸如 Alta Vista, Google 等各种搜索引擎更是为网络数据提供了强有力的工具。按网络数据内容的表现形式可

以分为数字型、文本型、声音型、图像和图形型、综合多媒体型。另外现在我们利用的清华大学的中国学术期刊网、万方数据期刊网、维普期刊网等,以及 CSSCI、CSTPC、SCI 等都提供网络版服务,为我们进行计量和评价提供着强有力的数据保障。

1.1.1.2 文献计量分析法的方法体系

文献计量分析法的方法体系主要由统计分析法、数学模型分析法、引文分析法、词频分析法、共现分析法、聚类分析法和计算机辅助信息计量分析法等方法组成。^①

(1) 统计分析法

这里所指的统计分析法包括两个部分,一是文献计量的一些专业统计术语。比如馆藏量、流通量、引文量、时差系数和词频分析等;二是指数理统计,主要包括随机样本、抽样分布、参数估计、假设检验、回归分析、方差分析和聚类分析等。概括起来,数理统计的研究主要是两个方面:一是研究如何从总体中抽出样本,抽多少,怎样抽的问题,即抽样方法问题;二是研究如何对抽样的结果(样本数据)进行合理的分析,作出科学的推断,即统计推断问题。数理统计在一定程度上是由于人们无法处理过多的数据而产生和发展起来的,适用于处理随机性的统计对象。

(2) 数学模型分析法

由于文献计量分析法具有定量学科的性质和担负着定量分析的任务,因此数学模型方法是文献计量分析法的最基本的方法之一,也是一种常用的有效的分析研究方法。数学模型是用某种数学语言描述系统的结构和行为而构造出来的,它除了从宏观的角度模拟出系统的结构和运动规律外,主要是从微观上、从量的方面描述出系统的结构、联系以及运动规律。因此,数学模型具有更为精细和准确等特点。数学模型按照不同的划分标准和方法可以划分为不同类型,比如按照表达的形式来划分分为三类:①解析式和图像模型。②方程组模型。③图表模型。建立数学模型的步骤基本为以下六步:①确定目标。②搜集原始数据。③建立理论模型。④确定参数。⑤验证理论模型。⑥预测和决策。数学是一门重要的基础学科,其必将在文献计量分析法的日后发展中起到越来越重要的作用。

(3) 引文分析法

所谓引文分析(Citation Analysis),就是利用各种数学及统计学的方

① 邱均平等. 信息计量学. 武汉: 武汉大学出版社, 2007

法和比较、归纳、抽象、概括等逻辑方法,对科学期刊、论文、著者等各种分析对象的引证与被引证现象进行分析,以便揭示其数量特征和内在规律的一种文献计量分析方法。几十年来,随着引文分析的广泛应用,其技术方法不断丰富,逐渐完善。从目前情况来看,引文分析大致有三种基本类型:①从引文数量上进行研究;②从引文间的网状关系或链状关系进行研究(如共现分析);③从引文反映出的主题相关性方面进行研究。当然若从引文的其他不同特征出发,还可派生出许多类型的引文分析。例如,从引文的语种、国别、年代、作者等进行的引文分析。引文分析作为文献计量分析法自身独特的方法,得到了很大的关注,应用也越来越广泛,尤其在信息检索和信息预测等方面有着广阔的前途。

(4) 计算机辅助信息计量分析法

随着计算机的日益普及和信息技术的推广应用,我国文献计量分析法的 research 手段逐步从手工统计分析向计算机辅助的统计分析过渡,出现了明显的自动化趋势。20 世纪 90 年代以来,对计算机辅助的文献计量分析的研究不断增加,取得了不少成果,其应用也越来越广泛。计算机辅助文献计量研究在理论分析的基础上,着重进行了计算机辅助文献计量软件的设计与开发工作,实现了套录资料重组建库以及多种资料统计分析功能;计算机辅助计量分析方法的建立和成熟,标志着我国文献计量研究的方法体系已基本形成,并日趋完善。特别是我国科技论文与引文数据库、中国科学引文数据库、中文社会科学引文索引数据库等一批数据库光盘或网络版的研究成功和出版发行,为文献信息的自动化统计和分析提供了必要的工具和条件,必将大大促进自动化趋势的进一步加强。

1.1.2 文献计量分析法研究进展

1.1.2.1 文献计量分析法的发展历程与现状

自从 1917 年动物学教授科尔(F. T. Cole)和博物馆馆长伊尔斯(N. B. Eales)对比较解剖学刊物的论文进行定量统计研究以来,起源于目录学研究的文献计量分析法从无到有、从小到大,获得了迅速的发展。随着洛特卡定律、布拉德福定律、齐普夫定律的创建,取得了以文献增长与老化规律、三大定律、引文分析为内容的主要研究成果。1969 年,英国学者普里查德(A. Pritchard)在其发表的《是统计书目学还是文献计量学?》论文中首先提出文献计量学(Bibliometrics)这一术语,并将文献计量学定义

为:“把数学和统计学用于图书和其他文字通讯载体的科学。”^①文献计量学研究主要集中在方法论(Methodology)、科学信息(Scientific information)、科学政策(science policy)三个领域,方法论也就是文献计量分析法是被公认的文献计量学的基本研究领域,主要用于图书情报服务中。^②随着文献计量分析法方法体系的逐渐成熟,该方法被广泛应用到众多学科文献信息的定量分析与研究中,并在实践的应用过程中不断得到新的拓展与延伸,促进了科学计量法、情报计量法与网络信息计量法的产生与发展。

洛特卡定律是1926年,由美国的统计学家、情报学家洛特卡(A. J. Lotka)在其发表的《科学生产率的频率分布》论文中提出的,在文中洛特卡统计分析了化学和物理学两大学科中一段时间内科学家们的著述情况,提出了定量描述科学生产率的平方反比分布规律,也被称为“倒平方定律”;布拉德福定律是1934年,由英国文献学家布拉德福(S. C. Bradford)提出的,该定律定量描述科学论文在相关期刊中集中——分散状况,经过后来的布鲁克斯、维克利、莱姆库勒等许多研究者的修正和研究,发展成为著名的文献分布理论;齐普夫定律是由美国语言学家齐普夫(G. K. Zipf)于1935提出的,经过朱斯、芒代尔布罗等的修正,形成了由齐普夫第一定律和齐普夫第二定律构成的文献词频分布的规律。1950年普赖斯在1944年赖德(A. F. Ryder)的发现的基础上提出文献呈指数增长的方程,1958年,美国学者贝尔纳(J. D. Bernal)将“半衰期”概念的引入和1971年“普赖斯指数”的提出,使文献老化的测量与文献老化定律的研究取得了进展,巴尔顿—凯普勒方程描述了文献老化的规律。1963年《科学引文索引》(Science Citation Index, SCI)的创刊,支撑了引文分析研究的深入进展,引文分析方法的日渐成熟使文献计量分析法的研究进入到了一个新的发展阶段。

随着几个定律的诞生,国内外众多学者与实践者,一方面主要依据经验方法、理论方法和模型方法,对各个定律学科的选择、数据的收集、参数取值范围的确定、方程的拟合及检验进行了检验性与拓展性研究,试图在原有定律的基础上探索新的现象和形成新的方法。其中,20世纪60年代开始的布—齐—洛三定律一致性研究引起了世人的关注,并取得了重大的成就。肯德尔(Kendall)提出布拉德福定律与齐普夫定律一致性后,

① 罗式胜. 文献计量学概论. 广州:中山大学出版社,1994

② W. GLÄNZEL. Bibliometrics as a Research Field; A course on theory and application of bibliometric indicators. COURSE HANDOUTS, 2003

费尔桑则进行了系统性研究,之后,休伯特(Hubert)和耶伯郎斯基(Yablonsky)对三个定律进行的一致性全面探讨,以及普利亚(Polya)的单缸模型和泰格(Tague)多缸模型,不仅很好的揭示了布—齐—洛三定律分布的机理,而且提出了描述分布状态的统一的数学模型,这是都为后续的实践应用奠定了方法论基础。另一方面,将已有的定律用于实践当中以解决面临的问题或者得出有价值的结论。如在洛特卡定律的研究与应用方面,戴维、利芬斯、西蒙、普赖斯、肖尔、维拉奇、欧帕、格勒塞、Y. S. Chen等众多科学家对洛特卡定律进行了深入研究。20世纪60年代以后掀起了研究的高潮,70年代产生了以维拉奇为代表的系统性研究成果并提出了广义的洛特卡定律,80年代以欧帕为代表的科学家在洛特卡定律参数估值上提出了改进的新方法,1987年古柏塔对昆虫学文献作者分布的研究使人们真正关注起合作者问题,并依据普赖斯思想提出了合作者分布模型,^①1997年,Rousseau等人对拟合如洛特卡定律一类的指数分布定律软件的开发,使洛特卡定律在网络环境下的应用得到了拓展。^②

1.1.1.2.2 我国文献计量分析法的发展趋势^③

(1) 科学化趋势

1969年,英国著名情报学家阿伦·普里查德(Alan Britchard)正式提出了“文献计量分析法”(Bibliometrics)这一术语,以取代当时已有47年历史但却很少使用的“统计书目学”(Statistical Bibliography)名称。这一术语的提出标志着文献计量分析法的正式诞生,真正开始从学科的角度来研究文献计量问题。在其后的30多年中,国际文献计量分析法获得了迅速发展,已经成为图书情报领域最活跃的一个分支学科,代表着情报科学的主流方向和学科发展趋势。在这种科学化趋势的影响下,在我国文献计量分析法研究和发展中,我们也特别强调要提高文献计量分析法的科学性和精确性,要确立文献计量分析法的学科地位。因此,我们着力从学科建设的角度来论述文献计量分析法的学科内容,从理论、方法和应用三个方面构建了文献计量分析法的学科结构体系。当前,文献计量分析法发展的科学化趋势主要表现在以下三个方面:一是要继续深化和完善

^① 徐久龄,刘春茂,张辉. 洛特卡定律的研究. 见:情报学进展 2000-2001 年度评论(第四卷). 北京:国防工业出版社,2001:1-39

^② Brendan Rousseau, Ronald Rousseau. LOTKA: A program to fit a power law distribution to observed frequency data. Cybermetrics, 2000(1):4. <http://www.cindoc.csic.es/cybermetrics/articles/v4i1p4.html>

^③ 王宏鑫,邱均平. 21 世纪文献计量学的发展趋势. 高校图书馆工作, 2000(4)