

二次情报资源的开发与利用研究*

——《中文自然科学引文索引》的综合开发与统计分析

邱 均 平

在图书情报领域内,情报资源既包括原始的一次情报,也包括二次情报和三次情报资源。从广义上说,情报资源的开发与利用研究也应当是各类型各层次情报资源的全方位的研究,其内容范围十分广泛。作为二次情报的一种特殊类型的引文索引的综合开发,自然是整个情报资源开发与利用研究的重要内容之一。但是,从目前国内的有关研究来看,大多数研究往往停留在一次情报资源上,而不恰当地忽视了二次情报资源的全面开发,特别是从不同角度,对业已编印出来的中文科学引文索引的多种功能进行综合性开发和利用的研究更是没有开展起来。因此,笔者选择《中文自然科学引文索引》(以下简称《索引》)作为具体对象,对其进行较全面的统计和定量分析,比较系统地探讨引文索引的应用开发和引文规律问题,以推动我国二次情报资源的综合开发与利用研究的深入发展。

一、引文索引综合开发与利用研究的意义

引文索引是本世纪60年代才诞生和发展起来的一类新型索引,其中最早问世和最杰出的代表当推美国科学情报研究所(ISI)编辑出版的《科学引文索引》(SCI)。这种引文索引系统引起了国际科学界的极大关注和广泛利用,其发展之快、影响之大是十分惊人的!究其原因,除了引文索引本身具有多种特殊功能、顺应科学发展的客观需要等因素外,从一开始就相当重视其开发与利用研究,广泛开展宣传、应用和推广工作,不能不说是其获得巨大成功的一个重要原因。该索引的创始人E·加菲尔德就曾率先发表许多文章,反复介绍引文索引的原理和编制方法、宣传其功能、应用领域和利用技巧等。目前,国外每年发表的分析、评价与应用引文索引的文章就达几千篇之多。我们应当及时分析和借鉴国外的成功经验,在研究和编制中文科学引文索引的同时,不失时机地把对引文索引的开发研究和推广应用工作提到议事日程上来;采取“边编制、边开发,以编制为基础,以开发促编制”的发展方针和策略,以利于我国科学引文索引建设的顺利发展。具体来说,综合开发引文索引的必要性和重要性可从以下几个方面来认识:

1. 有利于普及和完善引文分析方法。在一定意义上说,引文分析方法是在引文索引的基础上发展起来的。引文索引的出现,不仅为引文分析方法提供了一种较为系统的和有效的工具,而且其开发和应用的程度体现着引文分析方法不断普及和完善的过程。

2. 有利于我国文献计量学向应用阶段发展。文献计量学由纯粹理论研究向实际应用阶段发展,一方面是由学科的内在规律决定的,另一方面也有赖于大型的文献计量工具的开发利用。国外文献计量学已进入广阔的应用阶段,而实现这种过渡的一个重要条件就是美国《科学引文索引》的研制成功和出版发行。开展我国文献计量工具和手段现代化的研究是我们面临的一项十分紧迫的任务,我们必须及时编制和开发中文科学引文索引,以满足我国文献计量学研究的迫切需要。对中文科学引文索引的开发与利用,必然会促进我国文献计量学大规模的和系统的应用研究的开展。

3. 有利于发挥工具性刊物的综合效益。一般来说,引文索引的编制首先往往是以情报检索为其出发点的,但后来的发展在很大程度上取决于是否实现多功能的综合开发与利用。美国《科学引文索引》的发展历程

* 此文系武汉大学科研基金项目“科技情报资源的开发与利用研究”报告之一。

就是一个明显的例证。事实上,只是当它不仅用于情报检索,而作为文献计量和科学计量的有力工具广泛应用于引文分析、科学学、人才学、预测学等许多学科和工作领域时,才充分显示出强大的生命力。因此,我们从现在起就应当高度重视中文科学引文索引的多功能开发,以充分发挥其社会效益和经济效益,从而促进我国工具性刊物体系的建设和完善。

4. 有利于情报资源的开发与利用研究。对引文索引的开发和应用,无疑是整个情报资源开发与利用研究的重要组成部分之一。同时,引文索引把大量一次文献收集起来,组织成一个有序的方便利用的文献体系,从而有利于各类文献情报资源的开发与利用。

二、文献评价功能的开发与统计分析

从总体上说,引文索引的基本功能主要表现在两个方面,即引文分析功能和情报检索功能。由于引文分析的应用相当广泛,所以其功能也体现在许多方面,文献评价功能就是其中之一。

利用引文索引评价文献主要是基于:引文索引提供的文献被引证量在一定程度上反映文献的质量、学术影响及其在科学上所起的作用。一般来说,科学论文、期刊或专著等文献的被引次数越多,则说明该文献至少在某一方面具有较大价值和特色,其学术吸引力和作用越大。这种客观量度是通过引文指标来衡量的。例如,美国《科学引文索引》(SCI)和《期刊引证报告》(JCR)提供的被引次数、影响因子、当年被引指数等指标都可以用来定量地评价文献。对于《中文自然科学引文索引》来说,它有一种按被引作者姓氏笔画排列的“引证索引”,提供了被引证项(被引证文献作者姓名、刊名、卷[年]、期、页)和引证项(引证文献作者姓名、刊名、卷[年]、期、页、篇名)的特征信息。若对此进行分项统计,则可获得被引文献的有关特征信息数据,为进行文献评价提供定量依据。通过文献评价,可以测定关键论文、核心期刊、重点图书,以及进行各类型、各语种、各国文献的功能和利用率分析等。

为了获取数据和进行开发研究,笔者对《中文自然科学引文索引》(兰州大学版)第1—5分册进行了具体统计,涉及的引用时间为4年,被引期刊576种、被引作者4093个、被引总次数达9,189次。其数据量较大,具有一定的代表性和说服力。利用统计数据,这里仅作两个方面的评价与分析:

1. 期刊评价。从统计数据来看,被引论文次数按被引期刊的分布具有明显的集中性与离散性,也就是说存在着强烈的“核心期刊效应”。这种文献分布特征是进行期刊评价和选择的客观基础。在实际工作中,期刊评价的目的之一就是要测定核心期刊,而核心期刊的测定实质上就是期刊质量和价值的评价问题。测定核心期刊的关键是要选择和采用较为科学的、具有定量分析的合适方法,而引文分析法是最常用的有效方法之一。根据引文分析方法和实际统计数据,我们可以大致确定:在1980—1983年4年间被引频次较多的排在前面的109种期刊作为我国自然科学核心期刊。它们都是被引10次(含10次)以上的期刊。虽然其种数只占被引期刊总数的18.92%,而提供的被引次数却达到被引总数的83.77%。值得注意的是,在这些核心期刊中有几种社会科学期刊被列选,这是自然科学与社会科学相互沟通与渗透的反映。

2. 类型分析。对被引期刊的类型进行分析后可以发现,综合性期刊和学报类期刊的被引率较高,占有相当重要的地位。对15种(占被引期刊总数的2.6%)综合性期刊统计,被引次数达1,763次,占被引总数的19.2%;其中,仅《中国科学》和《科学通报》两种期刊就被引1,679次,占被引总数的18.3%。这一方面说明这些期刊具有较高的质量和较大的学术影响,另一方面也是由于它们涉及的学科较多,其读者面和使用面较广所致。统计数据还表明,学报类期刊所占比重较大,其被引期刊191种,只占总数的33.16%,而被引次数竟占总数的51.05%。而且在各学科的被引次数中占的比例也很高。例如,在天文学中占85.96%;在物理学中占78.30%。其中,大学学报被引112种,占总数的19.4%,被引720次,占总数的7.8%。这从一个侧面反映了科研机构人员与大学人员的联系。同时,在109种核心期刊中,学报类期刊就有56种,占到51.38%。这可能与《索引》收录的引文来源主要是学报类期刊有关。但这至少从一个侧面说明,对于我国自然科学研究来说,学报类期刊无疑是一类十分重要的情报源。

三、情报理论研究功能的开发与统计分析

情报学理论是以情报和情报工作规律为基础的。情报的产生、分布、增长、交流、传递、利用、老化等规律是情报学理论研究的基本内容和重点课题。而利用引文索引提供的丰富的数据资料,可以对情报流和情报工作规律进行许多方面的研究和探讨。

1. 文献情报产生和增长规律的研究。在科学著述活动中, 比较严谨的科学论文一般都附有参考文献或引用书目。所以, 已发表的绝大多数科学论文(收录范围内的)都会成为“引证文献”而被收入引文索引中。一个完整的引文索引系统不仅收集了每年涌现的大量文献情报源, 而且还提供了从引证作者的角度查找和统计文献生产信息的索引途径, 例如SCI中的“来源索引”。因此, 通过引证文献的统计, 可以掌握和了解文献情报生产的数量和特征, 从而分析情报产生的方式、环境、载体、时间, 以及各著者、团体、国家的情报生产能力等方面的规律。若对不同年份的引文索引进行统计和比较分析, 则可从数量上揭示整个科学或各学科、专业的文献情报的增长和变化规律。

2. 文献分布的集中与离散规律的研究。科学文献的集中与离散分布规律表现在许多方面, 具有一定的普遍性。利用引文索引及其统计数据, 可从不同角度对此进行揭示和验证。例如, 日本学者曾利用1980—1982年的SCI和SSCI(社会科学引文索引)中日本论文的引文数据, 验证了著名的科学文献分布定律—布拉德福定律的正确性。根据SCI数据库提供的资料, 美国情报学家加菲尔德(E·Garfield)曾提出“引文集中定律”。对《中文自然科学引文索引》的统计分析也得出了类似的结果, 发现我国的自然科学期刊引文按被引期刊的分布不仅符合加氏定律, 而且其集中化趋势更为明显。《索引》反映的中文自然科学期刊引文的分布情况如图1所示。

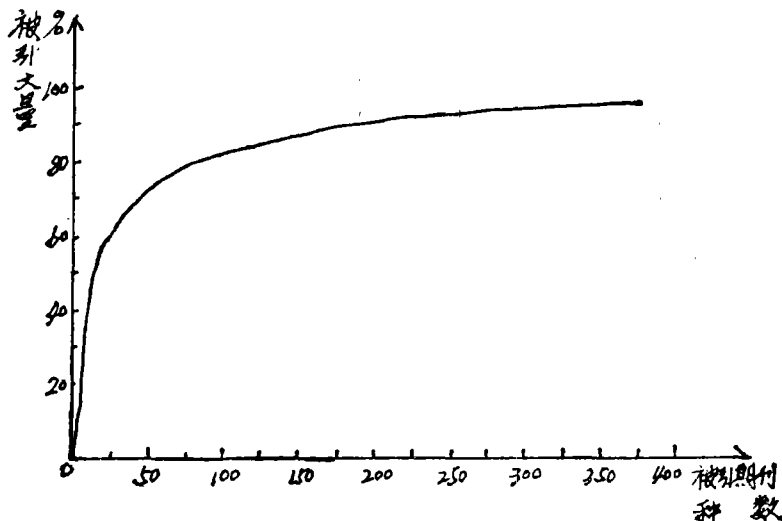


图1 中文自然科学期刊引文的分布

数据表明,《索引》收录的所有被引文献的一半仅来自12种期刊, 只占被引期刊总数的2.08%; 40种期刊(占被引期刊总数的6.94%)就提供了70%的被引文献; 被引总数的75%只涉及到55种期刊; 而约三分之一的被引期刊就承担了90%的被引次数; 另一方面, 10%的被引次数却高度分散在近70%的被引期刊中。这充分反映出被引文献分布的集中与离散特性。

3. 情报传递和交流规律的研究。苏联著名情报学家米哈依诺夫曾经指出:“科学情报从它的创造者传递到它的需要者, 主要不是通过他们的私人交往, 而是通过科学文献系统, 更确切地说是通过科学技术文献系统来实现的”。科学文献从发表到被引用正好体现着情报从创造者到需要者的传递过程; 文献的相互引证关系客观地反映了情报交流的情况。由于引文索引是建立在科学文献的引证与被引证关系的基础上, 其独特之处就是它既反映了文献的产生与存在, 又同时反映了文献的利用与吸收。利用引文索引提供的“引文网络”, 对文献情报流的引证与被引证信息进行统计分析, 便可揭示情报传递的方向、方式、程度和交流特性等。因此, 科学引文索引是研究科学文献系统的情报传递和交流规律的一种有效工具。

在文献计量学中, 对于文献情报交流的数量和程度, 可通过一系列定量指标来测度。一般来说, 科学文献的引证关系有两类测度指标, 一是从引证角度出发的“引证量”、“引证系数”(或称引证率)等; 二是从被引证角度出发的“被引证量”、“被引证系数”(或称“被引证率”)等。从广义上说, 相对于不同的文献主体对象, 有各种相应的引证率和被引证率指标, 如论文、期刊、著者、学科、语种、国家、学术机构和出版单位等的引

证率和被引证率。“文献交流比”则是衡量各文献之间相互引用的一个相对量度指标。针对不同的文献主体,亦可规定出相应的交流比概念,如“期刊交流比”、“学科文献交流比”、“A国与B国文献交流比”等。具体来说,A刊与B刊的交流比可以定义为:

$$\alpha_{\text{期刊}} = \frac{\text{A刊引用B刊的文献数量}}{\text{B刊引用A刊的文献数量}}$$

通过对《索引》中反映的期刊引用数据的统计,获得了有关期刊的互引数据。根据这些数据和定义公式计算的 α 期刊值列于表1中。从中可以看出,总数 α 值小于1的《中国科学》、《数学学报》、《物理学报》等期刊是“文献输出期刊”,说明被其他刊物利用较多,贡献较大;而 α 值大于1的《科学通报》、《化学学报》等期刊为“文献输入期刊”,相对来说被其他刊物引用不够。若两刊的 α 值为0,说明它们之间没有引文上的联系,或者只有单向引用。例如,《数学学报》除了与综合性期刊互引之外,与其它学科的学报都没有引用与被引用关系,且 α 值很小,说明其独立性很强,作为一种工具性学科的刊物被利用较多,而引用别刊较少。又如,同属于生物学的《植物学报》与《动物学报》也无联系,说明它们具有很强的专业性。这都在一定程度上反映了期刊之间情报交流的流向、程度和特点。

表1 自然科学几种主要期刊互引的 α 值

引 证 期 刊 被 引 证 期 刊	中国 科学	科学 通报	数学 学报	物理 学报	化学 学报	植物 学报	动物 学报	地质 学报	地理 学报	总数*
中国科学	1.0000	0.8333	1.8158	1.3585	0.5493	0.7143	0.7857	0.9333	0.6667	0.9153
科学通报	1.2000	1.0000	3.1364	0.8523	0.8219	0.7600	2.1667	1.6842	0.4138	1.1239
数学学报	0.5507	0.3188	1.0000							0.4348
物理学报	0.7361	1.1733		1.0000	0.3333					0.9119
化学学报	1.8205	1.2167		3.0000	1.0000	0.1379				1.2348
植物学报	1.1000	1.3158			7.2500	1.0000		2.500	1.2000	1.8182
动物学报	1.2727	0.4615					1.0000			0.8846
地质学报	1.0714	0.5938				0.4000		1.0000	0.3333	0.6250
地理学报	1.5000	2.4167				0.8333		3.0000	1.0000	1.7297
总 数*	1.0926	0.8898	2.3000	1.0966	0.8098	0.5500	1.1304	1.6000	0.5781	1.0000

* 总数不包括本刊自引量。

以期刊为单位的引证率和被引证率可分别按下列公式计算:

$$\text{A期刊引用B期刊的引证率 } R_{\text{期刊}} = \frac{\text{A期刊引证B期刊数量}}{\text{A期刊文献的总引证数量}} \times 100\%$$

$$\text{A期刊被B期刊引用的被引证率 } R'_{\text{期刊}} = \frac{\text{A期刊被B期刊引证数量}}{\text{A期刊文献的总被引数量}} \times 100\%$$

如果公式中的A期刊与B期刊是同一期刊的话,则分别称为该期刊的“自引证率”和“自被引率”。利用统计数据求出的我国自然科学主要期刊的引证率和被引证率数据分别列于表2和表3中。这些指标数据定量地反映了各期刊之间情报交流的范围、程度和相对重要性等。例如,《中国科学》、《科学通报》、《化学学报》的交流范围较广;而《中国科学》、《科学通报》、《数学学报》、《物理学报》之间的情报交流比较充分;对各科学报来说,C类期刊(综合性期刊)显得十分重要。

4. 文献利用规律的研究。科学引文是文献利用的客观记录之一。它既反映了科学作者利用的文献范围、

表 2

自然科学几种主要期刊的引证率(%)

被引证 引证期刊	中国 科学	科学 通报	数学 学报	物理 学报	化学 学报	植物 学报	动物 学报	地质 学报	地理 学报	其 他	合 计
中国科学	19.20	10.73	3.90	4.06	2.20	1.41	0.62	0.79	0.68	56.41	100
科学通报	10.16	18.49	3.08	3.34	2.67	0.85	0.58	1.43	0.53	58.87	100
数学学报	9.05	5.24	41.90							43.81	100
物理学报	3.91	6.49		65.49	0.29					23.82	100
化学学报	6.83	7.03		1.15	34.26	0.39	0.19	0.10		50.05	100
植物学报	4.19	3.00			3.48	37.05		0.60	0.72	50.96	100
动物学报	2.73	1.17					26.32		0.58	69.20	100
地质学报	3.18	4.02				0.42		18.86	0.85	72.67	100
地理学报	3.33	5.37				0.93		2.22	16.48	71.67	100

表 3

自然科学几种主要期刊的被引证率(%)

被引证 引证期刊	中国 科学	科学 通报	数学 学报	物理 学报	化学 学报	植物 学报	动物 学报	地质 学报	地理 学报	其 他
中国科学	41.87	21.91	21.97	6.88	7.99	6.87	6.83	9.15	9.52	20.57
科学通报	28.08	47.87	21.98	7.16	12.30	5.22	8.08	20.92	9.52	27.20
数学学报	4.68	2.54	56.05							3.78
物理学报	6.53	10.15		84.81	0.82					6.65
化学学报	8.74	8.42		1.15	72.95	1.10	1.21	0.65		10.71
植物学报	4.31	2.88			5.94	84.89		3.27	4.76	8.75
动物学报	1.72	0.69					83.85		2.38	7.31
地质学报	1.85	2.19				0.55		58.17	3.18	7.06
地理学报	2.22	3.35				1.37		7.84	70.64	7.97
合 计	100	100	100	100	100	100	100	100	100	100

数量、时间、类型等情况,又从文献被利用的角度表征着文献的质量和利用效率。因此,通过《索引》中揭示的文献引证与被引证关系可以定量地考察和研究文献利用方面的规律。例如,一篇文章的引文数量及其变化往往是随机的,但从宏观的角度来看,一系列论文或期刊的引文量及变化却表现出自己特有的统计学规律。在图2中,除1980年的数字偏低外,其他各年的引文量基本上是稳定的。这表明《科学通报》和9种自然科学期刊的引文量年度变化具有一定的相对稳定性。

5. 文献老化规律的研究。利用《索引》进行科学文献老化规律的研究一般是以被引文献的年代分布为基础的。根据《索引》统计的引文年代数据可得到中文自然科学主要期刊引文的年代分布曲线,如图3所示。对引文年代数据和分布曲线进行分析,可以确定半衰期、最大引文年限、文献最佳保存期等。例如,以此计算出中文自

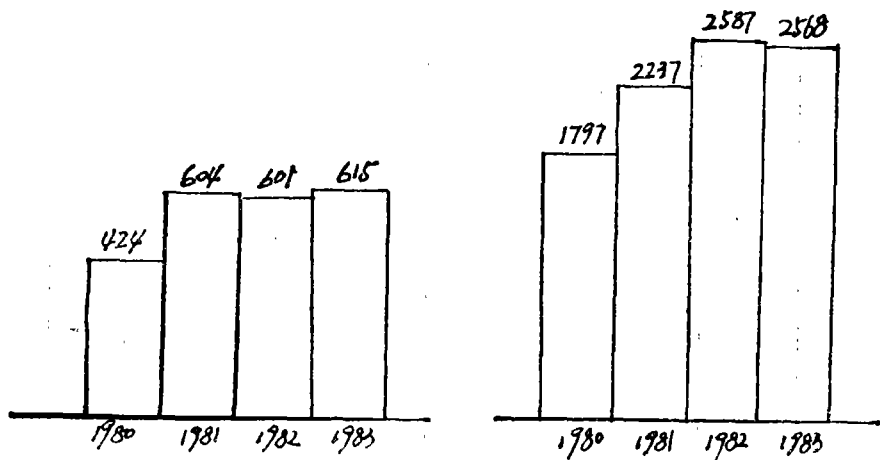


图2 《科学通报》的引文量变化

9种自然科学期刊的引文量变化

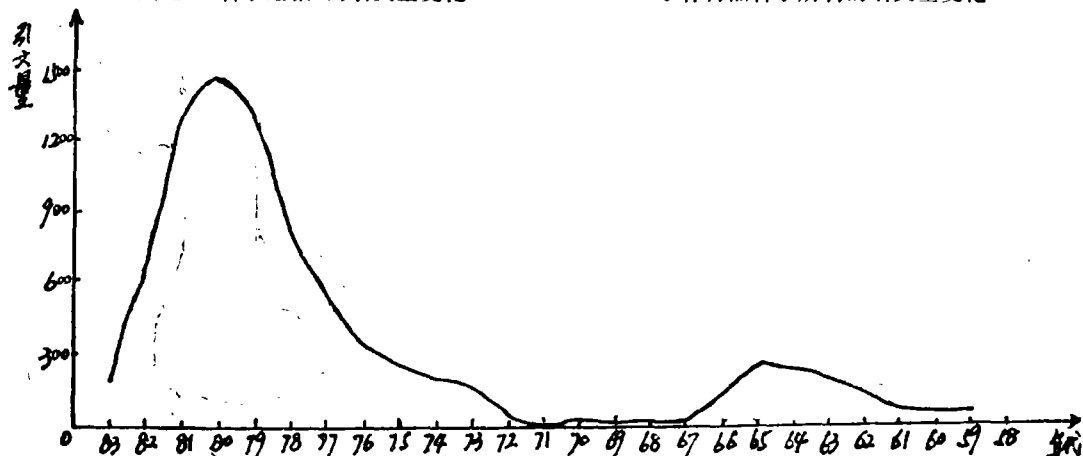


图3 中文自然科学期刊引文年代分布曲线

然科学期刊引文的半衰期为4.61年,从而反映出科学期刊文献的老化比整个科学文献的老化速度要快的特点。

四、人才评价功能的开发与统计分析

目前,国外普遍采用文献计量方法评价人材。特别是《科学引文索引》和《社会科学引文索引》的大规模出版,为文献计量学在人才评价方面的应用开拓了极为广阔的前景。利用文献计量方法评价人材的原理,主要是科学文献与科技人材之间的内在联系。具体来说,科技人员的成就和学术造诣不仅与其发表的文献数量有关,而且与其文献的被引数量有着一定的联系。科技人员发表的论文并不是每一篇都被引用为参考文献的。据国外统计,已发表的文献被引用10次以上的占25%,5次以上的占50%,2次的占9%,1次的占10%,约6%的文献却从未被人引用过。文献的被引次数越多,在一定程度上反映了文献的质量和价值,同时也能说明该文作者在学术界的影响和地位以及对社会的贡献。而且,对于某一著者来说,发表的文献越多,被引的次数也可能多。因此,著者文献被引率既从质量上也从数量上反映了文献的价值,因而可以衡量著者的学术水平。所以,科技人员发表的文献的被引率是评价科学人才的重要指标之一。而《索引》报道了大量的著者被引信息和数据,是开展人材评价的一种可供利用的便利工具。

统计表明,《索引》中反映的被引文献作者总数为4093个,被引9189次,平均每个作者在4年中被引2.25次;其中,个人作者3414人,占作者总数的83.41%,共被引8487次,占被引总数的92.36%;平均被引率为2.49次/人;而团体作者为679个,占16.59%;被引702次,只占总数的7.64%。相对于个人作者来说,团体作者者的被引率较低。具体的作者频次分布数据列于表4中。

表4

《索引》反映的被引文献的作者频次分布

作者被引率	被 引 作 者		被 引 数 量	
	人 数	占总数百分比	次 数	占总数百分比
1	1967	57.62	1967	23.18
2	604	17.69	1208	14.23
3	290	8.49	870	10.25
4	148	4.34	592	6.98
5	80	2.34	400	4.71
6	80	2.34	480	5.66
7	50	1.46	350	4.12
8	34	1.00	272	3.21
9	30	0.88	270	3.18
10	22	0.64	220	2.59
11—53	109	3.19	1858	21.89
合 计	3414	100	8487	100

从表4可见,分布频率最高的是4年中只被引过1次的作者,为1967人,占个人被引作者总数的57.62%;接着,被引2次、3次……的作者频率呈渐减趋势;被引10次(含10次)以下的作者共3305人,占个人被引作者总数的96.81%;共被引6629次占总数的78.11%。其具体分布趋势如图4所示。其中被引1次的作者频率与洛特卡定律中只发表过1篇文章的作者频率相吻合,其分布趋势也呈现出类似的规律性。同时,有少数的作者承担了较大量的被引次数。这部分被引率较高的作者可称之为“知名作者”。按照《索引》中出现的作者姓名统计的数据来看,在4年中被引10次以上的作者仅109人(排序名单从略),只占个人作者总数的3.19%,但提供被引数量1,858次,达到总数的21.89%;平均每人被引17.05次,是个人作者平均被引率的6.85倍。其中,被引率最高的一位名叫“范海福”的作者,4年中被引53次;第2名:郭常霖,被引48次;第3名:黄耀曾,40次;第4名:吴杭生,38次;第5名:唐敖庆,37次。从学科分布来看,物理学和化学领域知名作者的人数、被引次数及所占比例都较高,例如物理学中知名作者就有28人,占1/4多,平均每人被引20次以上;化学中有19人,平均被引率也达18.74次。在被引次数最高的前5名中,就有三人是物理学的,二人是化学方面的。这说明这些学科的发展比较成熟,在长期的研究过程中已形成一批著名的“科学精英”和研究十分活跃的学术带头人。

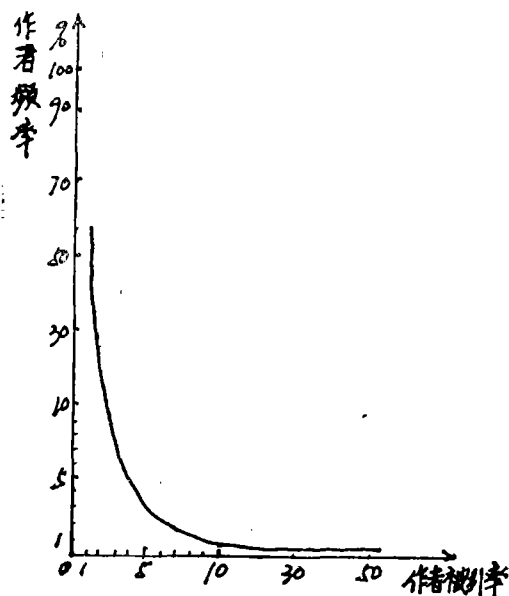


图4 被引作者频率分布

对我国自然科学的109位知名作者的被引情况进行分析可以发现,其被引率较高的原因主要是:①所发

表的文章富有创见,提出新的理论、观点或数据,或者开拓新的领域,对某一方面的科学研究具有指导意义或带头作用,也就是说“关键论文”的被引率较高;②作者“自引”起作用:109人中有自引的85人,占77.98%,自被引率为45.91%;③科研组或其成员对学术带头人、研究生对导师,或者学友之间的引用较多,例如,我国数学家张广厚与杨乐之间的文献互引比较频繁;④知名度高、学术影响大的科学家被引率高,而且被引涉及的时间较长,例如,国内外著名的地质学家李四光,虽然已去世多年,但其文献至今还常常被引用。然而,不管哪一种情况发生的引用,都可在一定程度上说明被引文献对引证文献的联系、作用与价值,从而反映被引作者的水平、贡献和学术影响力。利用作者数据,可以为制定科技人员的考核标准和管理方案、进行人才评价提供科学依据,并能进而探讨人才学规律。通过引文分析测度出来的人材评价结果又往往与科研成果数量、获奖、同行评议或社会评价等各种科学评价形式相验证。例如,唐敖庆教授是我国著名的化学家,而他的被引次数也名列前茅;其他许多著名学者,如竺可桢、叶笃正、汤佩松、宋健、周光召等都因被引率较高而被选入109位知名作者行列中。

五、科学研究功能的开发与统计分析

由于引文索引具有独特的结构,使之既揭示了文献的外部特征,又反映出有关文献在学科内容上的联系。这种由文献引证与被引证关系体现的学科相关性,使得大量科学文献以学科构成网络或“引文链”。利用引文索引提供的系统资料,对“引文网络”和“引文链”进行分析,便可找出文献间的学科联系,追溯学科知识的进化,从而探讨科学的横向的内在结构和纵向的发展变化规律。通过对《索引》的引文统计,这里作如下探讨:

1. 分析学科特点。《索引》中反映的各来源期刊引证的被引文献的学科分布数据体现出这些学科与外界联系的特点。例如,数学与其他学科引文上的联系较少,具有较强的学科独立性;物理学和化学与其他学科的联系较广,表现出较强的学科开放性;同时也反映其作者对不同学科的情报吸收和文献利用的广泛性。而地理学则表现出与社会科学联系较紧密的特点。

2. 研究学科结构及相互影响。这方面的研究可通过“学科文献交流比”、“学科引证率”、“学科被引证率”等指标来定量地进行。其概念分别定义如下:

A学科与B学科文献交流比 $\alpha_{\text{学科}} = \frac{\text{A学科引用B学科文献数量}}{\text{B学科引用A学科文献数量}}$

A学科引用B学科文献的引证率 $R_{\text{学科}} = \frac{\text{A学科引证B学科文献的数量}}{\text{A学科文献的总引证数量}} \times 100\%$

A学科被B学科引用的被引证率 $R'_{\text{学科}} = \frac{\text{A学科被B学科引证数量}}{\text{A学科文献的总被引数量}} \times 100\%$

根据统计数据计算得到的 $\alpha_{\text{学科}}$ 、 $R_{\text{学科}}$ 和 $R'_{\text{学科}}$ 的数值分别列于表5、表6和表7中。利用这些指标数据可以说明学科之间的结构联系、相对重要性和相互影响的程度等。例如,生物学与地学的交流比 α 值小于1,

表 5 自然科学几门主要学科互引的 α 值

被引证学科 引证学科	数 学	物理学	化 学	生物学	地 学	总 数*
数 学	1.0000					
物理学		1.0000	1.0741		4.0000	1.5517
化 学		0.9310	1.0000	0.8529	0.2500	0.8310
生物学			1.1724	1.0000	0.6000	0.6963
地 学		0.2500	4.0000	1.6667	1.0000	1.5714
总 数*		0.6444	1.2034	1.4362	0.6364	1.0000

* 总数中不包括自引量。

表 6

自然科学几门主要学科的引证率(%)

引证学科 \ 被引证学科	综合	数学	物理学	化学	生物学	地学	天文学	医、药	农、林	社会科学	其他	合计
数 学	40.00	59.52									0.48	100
物 理 学	14.82	0.15	81.05	2.14	0.44	0.59	0.22	0.37			0.22	100
化 学	19.92	0.10	2.60	54.67	2.79	0.19		5.97			13.76	100
生 物 学	12.32			2.53	63.33	4.45		10.91	5.64	0.37	0.45	100
地 学	15.22		0.20	0.79	9.88	63.64	0.10	0.29	2.67	6.52	0.69	100

表 7

自然科学几门主要学科的被引证率(%)

引证学科 \ 被引证学科	数 学	物理学	化 学	生物学	地 学
数 学	98.81				
物 理 学	0.79	97.43	4.54	0.61	1.12
化 学	0.40	2.39	88.89	2.93	0.28
生 物 学			5.32	86.34	8.40
地 学		0.18	1.25	10.12	90.20
合 计	100	100	100	100	100

表明生物学情报较多地流向地学领域,说明生物学对地学的作用较强,具有相对的重要性。同时,在地学的引证率分布和被引证率分布中,生物学都占有较大比例,也反映出这两个学科之间的联系比较紧密等特点。

3. 确定学科情报源的合理结构。学科引证率反映出某学科利用的情报源的分布结构,从而为组织专科情报、制定文献比例、合理分配经费等情报管理工作的科学化提供依据。例如,在生物学的引证率分布中,医药、农、林所占比例较大,其次是地学、化学等,因此,在组织生物情报源时,这些学科的文獻应占有相应的比例,以使其情报源的分布结构趋于合理。

此外,通过《索引》中反映的“引文链”或“引文网络”的变化,可以开展科技史的研究。利用引文索引数据和所反映的最新资料,还可论证科研项目、鉴定科技成果、评价学科发展,从而为科研经费管理、科学指标和科技政策的制定等提供定量依据。因此,《索引》是开展我国科学学特别是科学计量学研究的重要而有效的工具。

六、情报检索功能的开发研究

作为一种索引语言,引文索引无疑具有情报检索功能。与传统索引相比,它具有检索效率高、可进行“顺时序检索”、“滚雪球式”的循环检索,特别适用于多学科综合性课题的专题文献检索等特点。而且,对《索引》的情报检索功能的开发和利用研究成果还必将充实和发展情报检索的方法论体系。

(参考文献略)