

中文核心期刊 要目总览

2000年版

北京大学图书馆
北京高校图书馆期刊工作研究会

中文核心期刊要目总览

A Guide to the Core Journal of China

2000 年版
(第三版)

主 编 戴龙基 张其苏 蔡蓉华
副主编 武振江 史复洋等

北 京 大 学 出 版 社
北 京

图书在版编目(CIP)数据

中文核心期刊要目总览 2000 年版/戴龙基,张其苏,蔡蓉华主编. — 北京:北京大学出版社, 2000.6

ISBN 7-301-04597-2/G·593

I. 中… II. ①戴…②张…③蔡… III. 核心期刊, 中文-期刊目录 IV. Z87

责任编辑:胡双宝

封面设计:林胜利

书 名:中文核心期刊要目总览 2000 年版

著作责任者:戴龙基 张其苏 蔡蓉华

标准书号:ISBN 7-301-04597-2/G·593

出版者:北京大学出版社

地址:北京市海淀区中关村北京大学校内 100871

网址:<http://cbs.pku.edu.cn/cbs.htm>

电话:出版部 62752015 发行部 62754140 编辑室 62752028

电子信箱:zpup@pup.pku.edu.cn

排印者:中国印刷公司

经销者:新华书店

787 毫米×1092 毫米 16 开本 48.25 印张 1770 千字

2000 年 6 月第 1 版 2000 年 6 月第 1 次印刷

印 数:0001—5,000 册

定 价:190.00 元

A Guide to the Core Journals of China

(2000 Edition)

The Authors made bibliometric studies on over 10,000 Chinese journals published in the People's Republic of China, and they have identified the core journals in each field from these journals.

The method to identified core journals is as follows. First, for every Chinese journal over a period of three years (1995 – 1997), statistical countings of several numbers were performed, including the total number of articles presented by the journal, the number of articles abstracted or indexed by relevant abstract or index journals, the number of articles cited by other related journals in the field, the mean abstracted times of each article (abstracted rates), and the mean cited times of each article (i. e. impact factor). Secondly, a mathematical transfer was done to all the statistical data, then a weighted mean was calculated for these transferred data. Finally, journals in each field were arranged by the value as on the weighted mean. Those appeared on the top level of all ranks were selected as core journals in the corresponding field.

Totally, 75 ranking listings of core journals were created, with the coverage of all fields including philosophy, social science, humanities, natural science, medicine, agriculture and engineering. This book comprises all the listings with a classified arrangement based on the Library Classification of Chinese Libraries. Each identified core journal has a bibliographic description (the title and its English translation, the publisher, etc.) and a synopsis. There are 75 lists of professional journals attached to the Appendix, and there is a title index at the end of the book.

The Book has been highly evaluated by a committee of experts organized by State Education Commission of People's Republic of China.

This is the third edition of this book. The First and second editions were published in 1992 and 1996 respectively.

编辑人员

主 编 戴龙基 张其苏 蔡蓉华

副主编 武振江 史复洋

副主编 (以汉语拼音为序)

陈国棉	高 民	金碧辉	刘永春	刘元科	沈德来
武志宏	杨 毅	尹 源	张彩珠		

编 委 (以汉语拼音为序)

蔡蓉华	陈国棉	戴龙基	冯会芙	高 民	季 乔
贾凤琴	江 南	蒋 颖	金碧辉	李 莉	刘筱敏
潘敖生	潘海泉	卿 尉	沈德来	史复洋	苏玉华
王金玲	王 云	武振江	武志宏	杨翠兰	杨 毅
尹 源	苑玉兰	张彩珠	张春红	张海涛	张其苏

软件编制

陈文广	张彩珠	刘尚炎	漆 宏	范晓红	黄 颖
-----	-----	-----	-----	-----	-----

顾 问

(以汉语拼音为序)

- | | |
|-----|---------------------------|
| 白光武 | 中国科技信息研究所研究馆员 |
| 陈佳洱 | 国家自然科学基金委员会主任,中国科学院院士 |
| 承欢 | 清华大学图书馆研究馆员 |
| 迟惠生 | 北京大学常务副校长,教授 |
| 冯国健 | 科技部条件装备司综合处处长 |
| 姜晓辉 | 中国社会科学院文献信息中心数据库部主任,副研究馆员 |
| 阚延河 | 教育部社政司副司长 |
| 李惠国 | 中国社会科学院文献信息中心主任,研究员 |
| 李晓明 | 教育部高教司条件装备处处长,副研究员 |
| 林被甸 | 北京大学历史系教授 |
| 刘波 | 新闻出版署报刊管理司司长,高级记者 |
| 刘凤泰 | 教育部高教司副司长 |
| 刘桂林 | 清华大学图书馆馆长,教授 |
| 孟连生 | 中国科学院文献信息中心研究馆员 |
| 莫作钦 | 中国社会科学院文献信息中心研究馆员 |
| 潘龙法 | 中国学术期刊(光盘版)电子杂志社社长,教授 |
| 潘永祥 | 北京大学图书馆研究馆员 |
| 邱金利 | 中国人民大学书报资料中心主任,编审 |
| 沈玉兰 | 中国科技信息研究所业务处副处长,高级工程师 |
| 宋镇铃 | 中宣部出版局副局长 |
| 孙蓓欣 | 国家图书馆副馆长,研究馆员 |
| 万锦坤 | 中国学术期刊(光盘版)电子杂志社副总编,研究员 |
| 王明亮 | 中国学术期刊(光盘版)电子杂志社常务副社长 |
| 邬书林 | 中宣部出版局局长 |
| 武夷山 | 中国科技信息研究所副总工程师,研究馆员 |
| 徐金堃 | 国家自然科学基金委员会综合计划局局长,研究员 |
| 徐克敏 | 北京大学信息管理系教授 |

徐引麓	中国科学院文献信息中心主任,研究馆员
杨东梁	中国人民大学图书馆馆长,北京高校图工委常务副主任兼秘书长
袁名敦	北京师范大学信息技术与管理学系教授
郑国安	科技部条件装备司司长
周大挺	科技部条件装备司条件处处长
周和平	国家图书馆常务副馆长
朱 强	北京大学图书馆研究馆员,全国高校图书馆期刊研究会理事长
庄守经	北京大学图书馆研究馆员
祖广安	国家自然科学基金委员会出版中心副主任,研究员

中文核心期刊的文献计量学研究报告

本课题为国家教委人文社会科学研究‘九五’规划项目“核心期刊的文献计量学研究”的一个子课题,“核心期刊的文献计量学研究”由北京大学图书馆主持,该课题分为三个子课题:

1. 中文核心期刊的文献计量学研究——由北京大学图书馆及北京高校图书馆期刊工作研究会共同主持。

2. 国外人文社会科学核心期刊研究——由北京大学图书馆主持。

3. 国外科学技术核心期刊研究——由南京大学图书馆主持。

经过两年多的研究,已对核心期刊的筛选方法、核心期刊的界定、核心期刊表学科类目的划分、核心期刊筛选软件的设计与使用、核心期刊的作用与影响等问题作了大量的研究。并且将研究成果——筛选出的核心期刊表以工具书的形式出版。它们分别为:《中文核心期刊要目总览》2000年版(即第三版)、《国外人文社会科学核心期刊总览》2000年版及《国外科学技术核心期刊总览》2000年版。

现将子课题中文核心期刊的研究工作逐项介绍如下:

一、筛选核心期刊的理论依据

虽然期刊文献数量浩如烟海,但是它们在期刊中的分布,它们的流通与使用都存在一定的规律,那便是“集中与分散”规律,这些规律构成核心期刊的基础。

1. 一次文献在期刊中的分布规律

英国文献计量学家布拉德福^[1]首先发现了论文在期刊中的分布规律。他认为:“对某一主题而言,将科学期刊按其登载相关论文减少的顺序排列时,都可以划出对该主题最有贡献的核心区,以及含有论文数量与之相等的几个区。这时核心区与相继各区的期刊数量成 $1:a:a^2\cdots$ 的关系。”当时,布氏研究了地球物理及润滑两个学科,他取分区数为3,测得 $a=5$ 。这表明 $1/3$ 的论文刊载在3.2%的少数期刊上,而其余 $2/3$ 的文章则分散在大量期刊上。这少数信息密度大、载文量多的期刊可称为核心期刊。对不同的学科,尽管 a 值不同,但分布规律是相同的。布拉德福还给出图像描述法,认为:“半对数载文量-期刊数”图线的起始弯曲部分,对应于核心期刊。此后各国学者发现布氏的两种描述存在矛盾,纷纷以实测数据进行修正,分别发展了以这两种模型为基础的数学模式。

2. 二次文献的分布规律

二次文献指文摘、题录、索引等。它们所反映的期刊论文的分布,也表现出类似的集中与分散的规律。联合国教科文组织的一个委员会在1967年的一篇研究文章中写道^[2],“从物理学和化学领域的重要文摘杂志中发现了一条规律,它们所列出的或编成文摘的75%的论文,仅来自它们所收摘的全部期刊的10%。”这告诉我们,在对科学出版物进行组织时,也存在一个核心区,对应一组核心期刊。

3. 引文分布规律

某学科的有关论文中,所有被引用的文章在期刊上是怎么分布的?《科学引文索引》(SCI)的创始人加菲尔德回答了这个问题^[3]。他在1971年统计了2000种期刊中的约1,000,000篇参考文献后发现,24%的被引频次高的文章出自25种期刊,50%出自152种期刊,75%出自767种期刊,而其余的被引文章则散布在数量大得多的期刊中。加菲尔德的工作说明,被引文章在期刊上的分布也有一个较为集中的核心区与广为分散的相关区。

综上所述,文献从它的出现、情报加工、使用三方面所呈现出的核心效应及由此派生出的其他因素(载文率、影响因子、即时被引率、半衰期、共引关系等)的集中效应,在核心期刊的研究中具有重要意义,是测定核心期刊的理论依据。

二、核心期刊的概念与作用

由上可知,我们可以找出布拉德福式的载文量多的核心期刊,二次文献收录量大的核心期刊,加菲尔德式的被引次数多的核心期刊。但是,这些核心期刊都是从某一角度反映期刊的特点,而没有综合反映人们所希望的核心期刊的各种优点。在本课题的前期研究中,我们曾经这样定义核心期刊:

某学科(或专业,或专题)的核心期刊,是指该学科所涉及的期刊中,刊载论文较多的(信息量较大的),论文学术水平较高的,并能反映本学科最新研究成果及本学科前沿研究状况和发展趋势的,较受该学科读者重视的期刊。

这样定义的核心期刊,只是一个理想概念。由于统计工具(检索性期刊或数据库)总是存在这样或那样的问题,筛选方法总是有这种或那种缺陷,因此实际筛选出来的核心期刊只能从总体上大致反映学科期刊的状况。排序不可能完全准确,还会有某些重要刊物被遗漏,更重要的是核心区与非核心区之间并无客观存在的界线。因此**核心与非核心只是一个相对的概念,任何过分夸大核心期刊的作用,不恰当地使用核心期刊的做法都是错误的。**

本课题的主要研究成果——**中文核心期刊表只是一种参考工具书**,这里要特别强调“参考”二字。当文摘刊物选择文献源、图书馆选购期刊和为读者导读、教师研究生查找资料和选择读物、科研管理人员进行研究成果评价工作时,都可以把相关学科的核心期刊表作为选择的“参考”。“参考”的意义在于根据各自的需要做增删修改,而不是一成不变的搬来使用。尤其在评定职称的问题上,一定要依据评定的专业范围、学术级别等具体情况自己定出适合于本单位的“重要期刊表”,而不应不加选择地搬用核心期刊表。不同级别、不同性质的专业人员都用同一个核心期刊表评定职称,显然也是不合理的。核心期刊表的价值在于它能面对有各种不同需求的不同层次的用户,而用户们“参考”核心期刊表,经过甄别后选定自己需用的期刊,才是正确使用核心期刊表的方法,才能使它真正产生社会效益。向社会提供一种“参考工具书”,这便是我们研究并筛选核心期刊的初衷。

三、核心期刊筛选方法的分类与比较研究

根据上述期刊文献的集中效应,中外学者曾采用过各种各样的不同方法来筛选核心期刊,因此必须从不同的角度来分类归纳这些方法,分析各种方法的优缺点,加以比较、发展,最终形成本课题的筛选方法。

方法的分类与比较。

1. 根据评价指标分为:

载文量法——又分绝对载文量法与相对载文量法。前者按刊载相关学科文章的多少来选核心期刊,后者按相对载文量(=某刊刊载相关学科的文章数/该刊同时期的发文总数)来筛选核心期刊。本法优点:直观、简便;缺点:由于仅以数量取胜,所以只能选出那些刊载某学科文献量大的期刊作为核心期刊,结果必然偏颇。

文摘量法——以文摘量与文摘率作为评价标准。前者按相关学科文章被摘的次数来选出此学科的核心期刊,后者则按平均每篇相关文章被摘的几率来选出核心期刊(见四)。该方法优点:按选定的工具书做统计,简便易行,一般来说,文章皆择优文摘,因此某刊被摘量大,不但反映了它在数量上的优势,也在一定程度上反映了该刊在质量上的优势。(若以不加选择全部收录的题录、索引等为工具,则相当于载文量统计。)缺点:统计数据的质量受文摘刊物质量的制约(如文献源选择是否恰当,文摘员水平的高低,源期刊缺期造成的漏摘等)。

引文分析法——有被引量、影响因子,即时被引系数等^[4]。优点:这类指标反映期刊在相关研究领域内的活跃程度,反映期刊的水平及质量。用这类指标可以筛选出高质量的学术期刊作为核心期刊。缺点:源期刊的数量及选择的恰当与否,引文本身的缺陷如不恰当的自引、否定引用等,都大大影响引文分析的可靠性。

流通量统计法——包括借阅次数、复印次数统计等。优点:反映期刊受到读者重视和使用的程度,在一定程度上反映期刊的水平。缺点:受读者群性质及收集数据方法的影响,收集数据难度较大。

2. 根据对原始数据统计的处理方法分为:

区域法——及由之变形的各种方法。布拉得福将期刊按其刊载某学科相关论文的多少按降序排列,再分成载文量相等的几个区,取第一区中的期刊为核心期刊。

百分比法——是区域法的变形,即将期刊按上法排好序后不必分区,取累计评价量(载文量)占总量的一个适当的百分比对应的期刊作为核心期刊。

图像法——布氏做累积载文量—期刊数对数曲线,取上凹部分与直线部分的交点对应的刊为核心期刊。在上述三种方法中,百分比法得到最广泛的应用,区域法与图像法只在以研究布拉得福定律为目的的工作中才使用。

3. 按选用指标的多少分为:

单指标法——如前(见三、1.)。优点是方法简便,缺点是单指标只反映期刊的一个特性,难以选出满意的核心期刊。

多指标综合筛选法——数据的综合处理又分多种方法,如:求逻辑和法^[5]、加权平均法^[6]、模糊数学法^[7]、层次分析法^[8]、主分量分析法^[9]等。优点:多个指标如果能合理地配合使用,可能得到较单指标更为满意的结果。缺点是统计量多,计算复杂、工作量大。

四、本课题选用的方法

根据上节的分析和《中文核心期刊要目总览》(以下简称《总览》)第一、第二版的实践,采用多指标综合筛选法。以1995—1997年国内出版的中文现刊作为统计研究对象(港、澳、台期刊暂不计入)。筛选方法如下:

第一步:采集数据

筛选指标:

1. 被引量统计——以索引类电子出版物为统计依据。
2. 被摘量统计——选专业性文摘做统计依据,首选电子版的文摘刊物,不得已时才选书本式文摘用手工统计。
3. 被引量统计——选用《中国科学引文索引》数据库作统计依据,人文及社会科学、医药卫生、农业科学和工程技术中没有引文数据的部分,分学科选择来源刊,用手工统计被引量。
4. 载文量统计——以不加选择见文即录的电子版出版物做统计依据,少量未被收录的重要期刊靠手工统计补入。
5. 被摘率统计

$$\text{某刊的被摘率} = \frac{\text{此刊 1995-1997 年在某学科范围内被摘的次数}}{\text{此刊 1995-1997 年刊载此学科论文的篇数}} \quad (1)$$

6. 影响因子

在指定学科范围内 A 刊 1997 年的影响因子以 $IF^{97}(A)$ 表示

$$IF^{97}(A) = \frac{\text{1997 年“来源刊”对 A 刊 1995-1996 年文章的引用数}}{\text{A 刊 1995-1996 年可供引用的文章数}} \quad (2)$$

三年平均的影响因子^[10]比较复杂,要统计超出 1995-1997 年范围以外的载文量。故只采用 1 年的影响因子。

第二步:综合筛选

1. 初选结果构成矩阵

假设共有 I 种刊进入综合筛选,评价指标为 J 个(这里 $J=6$),得到原始统计数据所构成的矩阵

$$V = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1j} & \cdots & v_{1J} \\ v_{21} & v_{22} & \cdots & v_{2j} & \cdots & v_{2J} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ v_{i1} & v_{i2} & \cdots & v_{ij} & \cdots & v_{iJ} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ v_{I1} & v_{I2} & \cdots & v_{Ij} & \cdots & v_{IJ} \end{bmatrix} \quad (3)$$

其中 $i=1,2,3,\cdots,I$ (进入综合筛选的刊的编号,与排序无关); $j=1,2,3,\cdots,J$ (筛选指标的编号,在此 $J=6$,例如 $j=1$ 表示被引量, $j=2$ 表示被摘量……)

v_{ij} 表示第 i 个刊对第 j 个指标的统计量,例如 $i=3, j=2, v_{32}$ 是第 3 个刊的被摘量。

2. 求隶属度,构成评价矩阵

首先定义单指标隶属度:

第 i 个刊对第 j 个指标的隶属度定义为:

隶属度 1

$$c_{ij} = \frac{v_{ij}}{\bigvee_{i=1}^I (v_{ij})} \quad (4)$$

隶属度 2

$$c_{ij} = \frac{v_{ij} - \bigwedge_{i=1}^I (v_{ij})}{\bigvee_{i=1}^I (v_{ij}) - \bigwedge_{i=1}^I (v_{ij})} \quad (5)$$

其中 $\bigwedge_{i=1}^I (v_{ij})$ 为所有进入综合筛选的刊对 j 指标统计值中最小的一个, 同样 $\bigvee_{i=1}^I (v_{ij})$ 是取 j 指标所有统计值中最大的一个。

例如, $j=1$, 是载文量, 设统计结果中, 载文量的最大值为 328, 最小值为 1 (1 个完整的统计, 最小值都是 1), 第 i 个刊的载文量为 205, 则 $\bigvee_{i=1}^I (v_{ij}) = 328$, $\bigwedge_{i=1}^I (v_{ij}) = 1$, $v_{ij} = 205$, 隶属度 1: $c_{ij} = \frac{205}{328} = 0.63$, 隶属度 2: $c_{ij} = \frac{205-1}{328-1} = 0.62$ 。显然具有最大载文量的那个刊的隶属度为 1, 而具有最小载文量的那个刊的隶属度分别为 $\frac{1}{\bigvee_{i=1}^I (v_{ij})}$ 或 0。

其他刊的隶属度都在 0—1 之间。隶属度越大, 表示这个刊进入核心区的可能性越大。

隶属度 1 与隶属度 2 是从模糊数学的多种综合评判数学模型中^[11]挑选出来的适合于做核心期刊评价的两种模型, 二者都编入计算软件 (本书第二版就已提供此两种模型, 只是前言中未写出第一种模型的表达式), 供研究者任选一种使用。

将式 (3) 中的统计值 $v_{ij} = (i=1 \cdots I, j=1 \cdots J)$ 按 (4) 或 (5) 式逐个换算为 c_{ij} , 得到由隶属度构成的评价矩阵 C

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1j} & \cdots & c_{1J} \\ c_{21} & c_{22} & \cdots & c_{2j} & \cdots & c_{2J} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{i1} & c_{i2} & \cdots & c_{ij} & \cdots & c_{iJ} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{I1} & c_{I2} & \cdots & c_{Ij} & \cdots & c_{IJ} \end{bmatrix} \quad (6)$$

3. 加权平均

各指标的权重构成向量 B

$$B = (b_1 b_2 \cdots b_j \cdots b_J) \quad \sum_{j=1}^J b_j = 1 \quad (7)$$

以专家评定法选取权重。

对评价矩阵 C 做加权平均

$$A = B * C^T = (b_1 \ b_2 \cdots b_j \cdots b_J) \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1j} & \cdots & c_{1J} \\ c_{21} & c_{22} & \cdots & c_{2j} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{i1} & c_{i2} & \cdots & c_{ij} & \cdots & c_{iJ} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{I1} & c_{I2} & \cdots & c_{Ij} & \cdots & c_{IJ} \end{bmatrix}^T = (a_1 a_2 \cdots a_i \cdots a_I) \quad (8)$$

$$\text{其中 } a_i = \sum_{j=1}^J b_j \cdot c_{ij}$$

a_i 为第 i 个刊的隶属度,把刊依 a 值的大小排序,并且截取略少于上一版的刊作为核心期刊。

第三步:专家评审

将筛选方法、统计数据及筛选结果送专家评审,请学科专家鉴定筛选结果是否符合该学科的实际情况。根据专家意见作个别调整,即产生该学科的核心期刊表。

五、核心期刊的数量界定

1. 问题的提出

综合筛选完成后,期刊按综合隶属度降序排队,这时又出现一个难题,那就是应该取多少种期刊作为核心期刊。有一些同行专家认为《中文核心期刊要目总览》第一、二版所选的核心期刊过多。理由是:“根据布拉德福定律,核心期刊与专业性期刊的数量之比例应该控制在 3%~4% 左右”。而本书第一版的比例最大的类超过 60% (F1 世界经济类高达 62%)^[12]

上述说法是对布拉德福定律的误解,布氏两组知名的统计数据,其一为从 326 种期刊中选出 9 种应用地球物理学核心期刊(数据见表 1),又从 164 种期刊中选出 8 种润滑学核心期刊,核心期刊与参选期刊之比前者为 2.8%,后者为 4.9%,并非 3%~4%。布氏定律是一个经验定律,其价值在于它揭示了某一专业的文献在期刊中分布的规律,布氏仅仅提出了两组统计数据,这两组数据中的具体数字并没有经过大量实践的反复验证,因此这些数字只能作为后人研究时的参考,而不应该将它当成一个标准来引用,因为这两组数据给出的核心期刊与参选期刊之比就各不相同,4.9% 几乎是 2.8% 的两倍,而且这个 3%~4%,是核心期刊与“参选期刊”数之比,而决不是核心期刊与“专业期刊”数之比。什么是“参选期刊”,凡载有该学科文献的期刊,哪怕在统计时段(1928~1931,1931~1933)中只有 1 篇相关的文章,也是参选期刊。但是只有全部或大部分内容都属于该学科的才是“专业期刊”,对此二组数做一个最粗略的估计,比如认为平均每年至少发表 3 篇相关文章的刊才算该学科的“专业期刊”,则应用地球物理学有 20 种这样的期刊,其中 9 种被选为核心期刊,二者之比是 9:20=45%,润滑学则有 8 种平均每年发表 3 篇以上相关文献的刊,8 种刊都选作核心期刊,核心期刊与专业期刊的比为 8:8=100%,因此如果要引用布拉德福的数据,那么结论是《中文核心期刊要目总览》第一或第二版中任何一个学科的核心期刊与专业刊之比都小于布拉德福数据的比例,《总览》所选的核心期刊不是太多,反而是太少了。显然核心期刊与专业期刊之比为 3%~4% 的论据是错误的,45%~100% 的论据是未经反复证实的,因此它们都不能作为判断核心期刊多少的准则。那么截取核心区的原则是什么?核心区与非核心区有没有客观存在的界线呢?

2. 如何截取核心区——中外学者界定核心期刊的方法

如前所述(见一、1.)英国文献计量学家布拉德福首先提出“核心区”的概念,并且给出了划定核心区的方法,即布氏分区中信息密度最高的第一区为“核心区”。后人自然而然地将“核心区”中的期刊称为“核心期刊”。布氏的分区数为“3”,并将这三个区分别称为“核心区”、“相关区”及“边远区”。显然,核心区内的期刊占总期刊数的

$1/(1+a+a^2)$, 载有占总量 $1/3$ 的文献。维克瑞称布氏的表述为“布拉德福文献分散定律”, 并且证明分区数是任意的^[13]。如果沿用第一区为核心区的概念, 分区数越多, 核心区越小, 核心期刊数就越少。

后来高夫曼找出划分的限度^[14], 他认为, 为了不使最后几个区出现期刊数等的现象, 每个区的论文数 L 应受到限制, 令 $L > E/2$ 。E 为载有一篇论文的相关期刊数。这样确定的分区数称为最大分区数, 以 m 表示。 $m = D/L$, D 为相关论文总数, 这时有最小的核心区和最少的核心期刊。

核心区划得大的有: 美国 D. T. 霍金斯用文摘法确定电化学核心期刊时, 取累计文摘量达 70% 的区域为核心区^[15]; 联合国教科文组织 (UNESCO) 将累计文摘量达到总文摘量 80% 的区域划为核心区^[16], 对应的期刊称为核心期刊。我国的研究者绝大多数选用 70% 或 80% 的累计评价量对应的区域为核心区, 这里姑且称之为“70% 法”和“80% 法”。

还有一些按布拉德福曲线的形状来划定核心区的方法。布氏曾将期刊按其刊载某学科论文的数量从大到小排序, 以论文的累积量 $R(n)$ 作纵坐标, 对应的期刊量的对数值 $\log n$ 作横坐标作图 (见图 1), 他发现此区前半截是上凹的曲线, 后半截是直线, 他把曲线与直线的连接点 C 视作分界点, C 点以前的区域称为核心区, 自然前 C 种期刊就是核心期刊。显然 C 点所对应的文献累积量不一定是总文献量的 $1/3$, 布氏也未从理论上说明自己划定的两种核心区的大小是否相等。

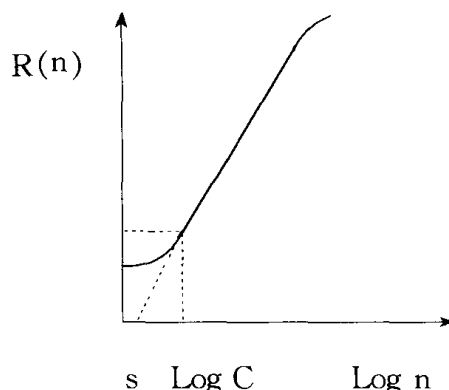


图 1 $R(n) - \log n$ 曲线

此外尚有三折线分割 $R(n) - n$ 曲线 (注意: 这个曲线的横坐标 n 不取对数, 见图 2) 的方法^[17], 三根直线的转折点将曲线分为三部分, 第一部分所对应的期刊即为核心期刊, 此法类似布氏的区域法, 但三个区域所对应的文献量一般不相等。用折线近似曲线, 在某种程度上反映了曲线的性质和形状, 但又并不像上一个方法选取曲线上有明显特征的 C 点作为分界点, 所以这个方法不同于以上的任何一种。

还有用正态分布特性来截取核心区的方法^[18]。

下面以布氏 1934 年使用的数据为例, 看看用前述各种方法可以划出多少种核心期刊。根据表 1 的数据, 用高夫曼方法决定最大分区数。 $E = 169$, 取 $L = 95 > 169/2$, 得最大分区数为 $m = D/L = 1332/95 = 14$, 最多可将表 1 中的期刊分为 14 个区域, 每区载有近似相等的论文数 (约 95 篇), 这时得到最小的核心区。由表 2 可以看出, 最少取 1 种核心期刊, 占总期刊数的 0.31%, 它刊载的相关论文数占总论文数的 6.98%; 最多取 108 种作核心期刊, 占总期刊数的 33.13%, 刊载 79.95% 的论文。从 1 种到 108 种之间, 还有许多不同的取法, 究竟哪些是合理的, 哪些不合理? 下面来讨论这个问题。

表 1:应用地球物理学期刊载文量排序表(1928 - 1931)

期刊数量	每刊载文 量	累计刊数 n	累计载文 量 R(n)	Log n	R(n) /N (%)	期刊数量	每刊载文 量	累计刊数 n	累计载文 量 R(n)	Log n	R(n) /N (%)
1	93	1	93	0.000	6.98	1	12	20	590	1.301	44.29
1	86	2	179	0.301	13.44	2	11	22	612	1.342	45.95
1	56	3	235	0.477	17.64	5	10	27	662	1.431	49.70
1	48	4	283	0.602	21.25	3	9	30	689	1.477	51.73
1	46	5	329	0.699	24.70	8	8	38	753	1.580	56.53
1	35	6	364	0.788	27.33	7	7	45	802	1.653	60.21
1	28	7	392	0.845	29.43	11	6	56	868	1.748	65.17
1	20	8	412	0.903	30.93	12	5	68	928	1.833	69.67
1	17	9	429	0.954	32.21	17	4	85	996	1.929	74.77
4	16	13	493	1.114	37.01	23	3	108	1065	2.033	79.95
1	15	14	508	1.146	38.14	49	2	157	1163	2.169	87.31
5	14	19	578	1.279	43.39	169	1	326	1332	2.513	100.00

注:末列 R(n)表示总载文量,R(n) /N 表示前 n 种刊载文量占总载文量的百分比。

表 2:用各种方法对表 1 数据截取核心期刊

方 法 (括号内的数表示每区所含的文献数)		核心刊数(种)	核心刊占总刊数的 百分比(%)	核心期刊载文量占总 文献量的百分比(%)
区 域 法	14 区(约 95 篇)	1	0.31	6.98
	7 区(190 篇)	2	0.61	13.44
	6 区(222 篇)	3	0.92	17.64
	4 区(333 篇)	5	1.53	24.70
	3 区(444 篇)	9	2.76	32.21
70%法		68	20.86	69.67
80%法		108	33.13	79.95
布拉德福图像法,C 点		9	2.76	32.21
三折线分区法		9	2.76	32.21

3.核心区与非核心区之间是否有明显的界线

如果回答是肯定的,那么只有用这个客观存在的界线来划分核心区与非核心区才是合理的,其他一切划法都不合理。但是,回答是否定的。因为用来评价期刊的文

献计量学指标,如载文量、文摘量等,它们在期刊中的分布都基本服从“布拉德福文献分散定律”。这些量在期刊中的分布曲线,例如 $R(n) - n$ 曲线(见图 2),是一根连续曲线,不存在任何突变点。

后人曾用多种函数形式来模拟这根曲线。函数连续的判据是它的一阶导函数存在,现以两种模拟函数来说明。

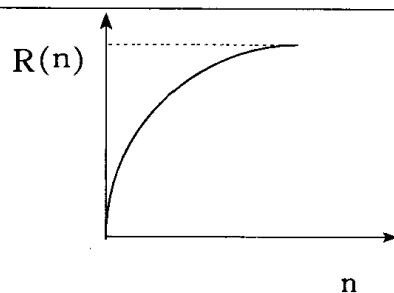


图 2 $R(n) - n$ 曲线

第一种,威尔金森公式^[19],它是有名的莱姆库勒公式的变形

$$R(n) = a \ln(1 + bn) \quad (9)$$

如前所述, $R(n)$ 表示前 n 种期刊的累积载文量。 a 与 b 是两个常数,此函数的一阶导数为

$$\frac{dR}{dn} = \frac{ab}{1 + bn} \quad (10)$$

在 $[0, \infty]$ 区间,即实半轴内有定义,即在期刊累积数的存在范围内一阶导函数存在,函数连续。

第二种,布鲁克斯公式^[20],描述 $R(n) - n$ 曲线,布拉德福将横轴取对数,如前所述,得到上凹部分与直线部分,形状与图 1 曲线类似,布鲁德斯用两个函数来描述它:

$$R(n) = \begin{cases} \alpha n^\beta & (1 \leq n \leq C) \\ k \ln \frac{n}{s} & (C \leq n \leq N) \end{cases} \quad (11)$$

其中 α, β, k, s 为常数, $R(n)$ 与 n 意义同上, C 是曲线两部分的交点, N 是期刊总数。

它的一阶导函数是:

$$\frac{dR}{dn} = \begin{cases} \alpha \beta n^{\beta-1} & (1 \leq n \leq C) \\ \frac{k}{n} & (C \leq n \leq N) \end{cases} \quad (12)$$

在 $[1, N]$ 区间上有定义,即一阶导函数存在,而且 $\alpha \beta n^{\beta-1}|_{n=C} = k \ln \frac{n}{s}|_{n=C}$ 及 $\alpha \beta n^{\beta-1}|_{n=C} = \frac{k}{n}|_{n=C}$, 函数连续。

如果要把图 1 的特点充分表现出来,可以做变量变换,令 $y = \ln n$, 则 $n = e^y$, 得到

$$Y(y) = \begin{cases} \alpha e^{\beta y} & (0 \leq y \leq \ln C) \\ k(y - \ln s) & (\ln C \leq y \leq \ln N) \end{cases} \quad (13)$$

$$\frac{dY}{dy} = \begin{cases} \alpha \beta e^{\beta y} & (0 \leq y \leq \ln C) \\ k & (\ln C \leq y \leq \ln N) \end{cases} \quad (14)$$

一阶导函数仍然在 $[0, \ln N]$ 区间内有定义,同样应有 $\alpha \beta e^{\beta y}|_{y=\ln C} = k(y - \ln s)|_{y=\ln C}$, 以及 $\alpha \beta e^{\beta y}|_{y=\ln C} = k$, 所以函数连续。证明不存在可以划分核心区与非核心区的突变点。

这里要特别讨论一个问题。本文第一部分曾提到用C点来划分核心区,C点虽不是突变点,但在半对数曲线中它是一个特殊的点,给我们划定核心区提供了依据。但是,正是C点的存在与否至今尚无定论,使我们无法用它作一个客观的分界点。众所周知,布拉德福定律的两种不同描述之间存在矛盾:

区域法 第1区刊数:第2区刊数:第3区刊数: $\cdots = 1:a:a^2:\cdots$ (15)

图像法 第1区刊数:第1,2区刊数:第1,2,3区刊数: $\cdots = 1:b:b^2:\cdots$ (16)

显然(15)、(16)两式中的分散常数 a 、 b 绝不相等,如果(16)成立,则半对数曲线有直线部分,如果(15)成立,可以证明此直线部分并不存在^[21],整条线都应该是弯曲的。布拉德福定律是一个经验定律,它的哪种叙述法更符合实际还在争论与研究之中。因此,结论是:

核心区与非核心区之间没有明显的、客观存在的界线,核心区是人为划定的,核心期刊数量的多少是任意的。本节第一部分列出的方法都没有错误,那么这些划法有无优劣之分呢?

4. 决定核心期刊数量的原则是什么

人们认为既称为“核心”,其数量就应该少;某些文献计量学家认为,核心期刊应该少于相关期刊总量的10%,并将之作为截取核心期刊的原则。此说虽然不错,却不全面,因为截取核心期刊的目的在于能够从较少的刊得到较大的累积评价量(如载文量、被摘量、被引量等,统称为评价量)。核心刊的少与评价量的大是相互制约的,而且对不同学科、同一学科的不同时段或同一学科用不同的评价指标筛选核心期刊,即使要求核心区对应同样的评价量百分比,所得到的核心期刊数占总刊数的百分比也是各不相同的,原因在于它们的分散程度不同。按布拉德福定律,将期刊按载文量(或别的评价量)从多到少排序,并划分为载文量相等的数个区域,其各区期刊之比为 $1:a:a^2:\cdots$ 。不同的研究对象其统计数据的 a 值——称布氏分散常数——都不相同。只要 $a > 1$ 就存在分散现象, a 越大,核心越集中,分散也越显著。下面通过计算展现不同 a 值产生的不同结果。

人们对布氏成功的三分区实验印象最为深刻,布氏对两个学科的数据进行分区,得到 $a_3 = 5$ (以 a 的下角标表示分区数, a_3 为三分区的分散系数)。再考虑两个极端的情况,设某个学科的 $a_3 = 2$,另一学科的 $a_3 = 10$,对这三种不同的分散情况求出文献量占30%~80%时核心期刊所占的比例。

莱姆库勒证明,对同一组统计数据作不同的分区时,各分散常数之间存在关系^[22]:

$$a_m^m = a_2^2 \quad (m = 2, 3, 4, \cdots) \quad (17)$$

由此可得 $a_3^3 = a_{10}^{10}$, $a_{10} = \sqrt[10]{a_3^3}$

当 a_3 分别为 2.0, 5.0, 10.0 时,有 $a_{10} = 1.231, 1.621, 1.995$ 。

第一个10%的文献对应期刊的百分比为 $a_{10}^0 / \sum_{i=0}^9 a_{10}^i$, 20%的文献对应百分比为 $\sum_{i=0}^1 a_{10}^i / \sum_{i=0}^9 a_{10}^i$, 以下依此类推。令 $A = \sum_{i=0}^9 a_{10}^i = 1 + a_{10} + a_{10}^2 + \cdots + a_{10}^9$, 得到下表: