

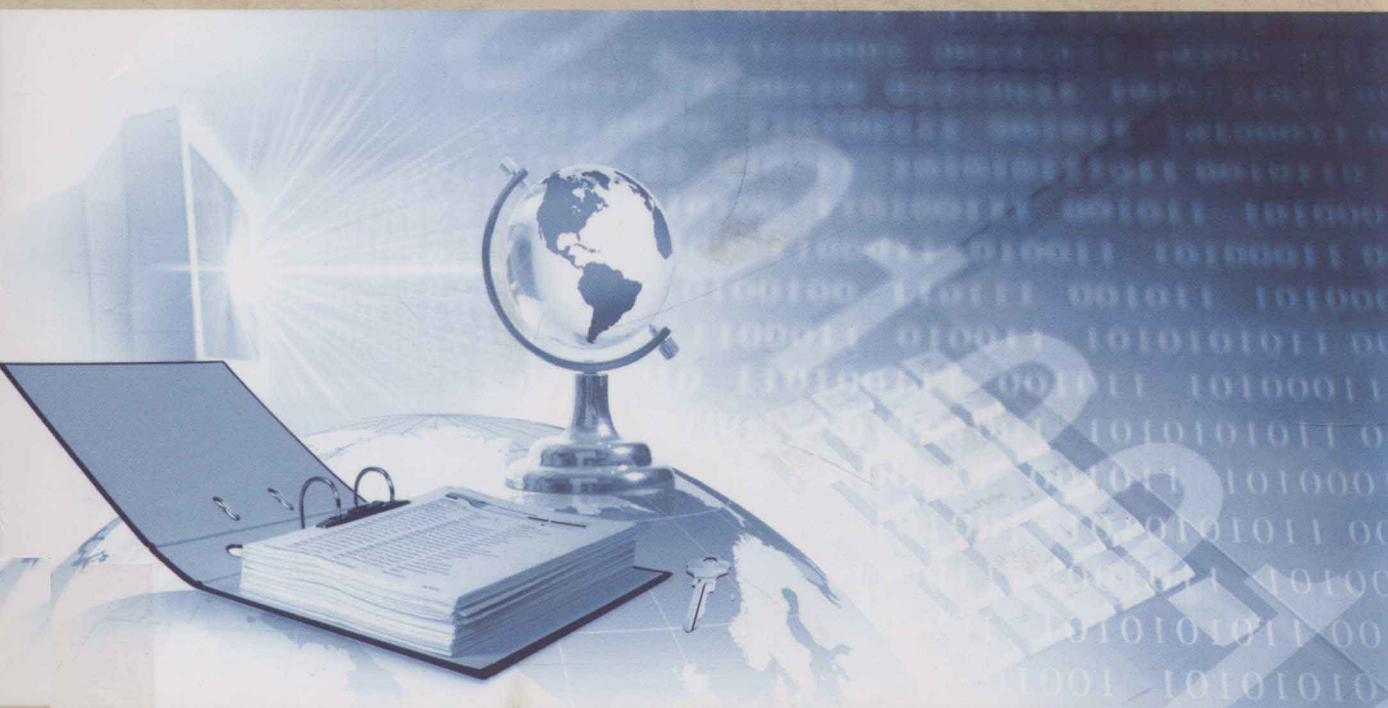


普通高等教育“十一五”国家级规划教材


情报学基础教程

(第二版)

叶 鹰 武夷山 主编



科学出版社



普通高等教育“十一五”国家级规划教材

情报学基础教程

(第二版)

叶 鹰 武夷山 主编

科学出版社

北京

内 容 简 介

本书是为学习情报学的大学生、研究生及科技情报、社科情报、图书情报工作人员编写的一部通用基础教材,内容覆盖情报理论与方法、情报管理与服务各相关领域,在统一的定量化情报学理论内核和方法基础上简明扼要地阐述了情报源、情报组织、情报检索、情报分析、情报技术、情报服务等情报学核心分支的基本知识和新近发展。

本书可以作为本科生和研究生的教材,也可供同等学力人员申请硕士学位者学习参考。

图书在版编目(CIP)数据

情报学基础教程 / 叶鹰,武夷山主编. —2 版. —北京:科学出版社,2012
普通高等教育“十一五”国家级规划教材

ISBN 978-7-03-035627-7

I. ①情… II. ①叶…②武… III. ①情报学-高等学校-教材
IV. ①G350

中国版本图书馆 CIP 数据核字(2012)第 224949 号

责任编辑:陈 亮 王京苏 / 责任校对:刘小梅

责任印制:阎 磊 / 封面设计:蓝正设计

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

新科印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2006 年 11 月第 一 版 开本:787×1092 1/16

2012 年 9 月第 二 版 印张:19 1/4

2012 年 9 月第一次印刷 字数:436 000

定价:38.00 元

(如有印装质量问题,我社负责调换)

编委会（按姓氏拼音顺序排列）

- 罗 勇 研究员，中国科学技术信息研究所研究生部副主任
- 马景娣 研究馆员，浙江大学图书馆副馆长
- 潘 卫 研究馆员，上海交通大学图书馆副馆长
- 潘有能 博士，浙江大学信息资源管理系副教授
- 王素芳 博士，浙江大学信息资源管理系讲师
- 武夷山 研究员，中国科学技术信息研究所副所长，南京大学信息管理学院
博士生导师
- 叶 鹰 博士，南京大学信息管理学院、浙江大学信息资源管理系教授、博
士生导师
- 张新民 博士，中国科学技术信息研究所研究员
- 张 旭 博士，研究员，中国科学技术信息研究所战略研究中心主任
- 曾建勋 博士，研究员，中国科学技术信息研究所信息资源中心主任

第二版前言

本书第一版于2006年作为普通高等教育“十一五”国家级规划教材出版后，受到情报学教学单位的欢迎，笔者在教学过程中也发现一些需要改进和优化的问题，于是借修订再版之机统一解决。

本次修订的总原则是加强理论联系实际、贯通学科基础与学科前沿。

为更好地理论联系实际，此次修订版的最大特点是邀请中国科学技术信息研究所的专家学者加盟撰写，参编者分别执笔各自熟悉的章节，以为本书增光添色。

各章的修订思路如下：

第1章情报学经验规律保留，旨在确立情报学之所以能成为科学的基础，内容以布拉德福定律、洛特卡定律、齐夫定律三大定态定律和文献增长律、文献老化律两类动态规律为主体，并依据国际前沿进展重写统一机理。

第2章是对情报学作为一门独立学科的理论解释，考虑到作为一门学科的教程必须有概念逻辑和体系结构，于是在介绍国内外代表性情报学基础理论的基础上，侧重对情报学理论体系的概念基础进行建构，在量化DIKW概念链的基础上构建了情报学理论体系。

第3章情报学方法论根据研究型教学的需要重写，侧重从宏观层面阐明情报学方法论体系，既为后续各章中的具体应用方法奠定统一的方法论基础，也为各章中微观层面的具体方法预留展开阐述的空间。

第4章情报源是情报工作和情报学研究的物质基础，修订版进行了改写和扩充。

从第5章到第9章，依次涉及情报组织、情报检索、情报分析、情报技术、情报服务，它们构成情报学各实体分支的主要内容，也构成情报学的核心课程体系，本次修订侧重简明扼要地阐明其知识内容和研究动向，以期为后续专门课程的学习奠定坚实基础。

情报教育作为情报学的专业教育领域，修订版另设新章阐述；当代情报学发展作为情报学的前沿，是研究型情报学教学的重要组成部分，故作为本书的最后章节。

各章在相对独立的同时保持了理论内核的严谨一致，并通过参考文献引导学生了解前沿研究。

希望第二版在保持公认知识内容并结合独特研究成果的同时，体现客观表述情报学理论知识并有机结合情报工作实践的创新，为普及和传播情报学知识贡献力量。

第一版前言

情报学作为一个独立的研究领域，是 20 世纪 50 年代开始确立的：50 年代初法语世界首先使用 L' informatique，1959 年英语世界正式提出 Information Science，从此，情报学逐步发展起来，至今仍处于动态发展之中。

如果在信息概念基础上认识情报，则情报是具有传递价值的信息；如果在知识概念基础上认识情报，则情报是具有传递价值的知识。因此，具有传递价值是情报的根本属性。也就是说，情报是针对特定对象的需要而提供的、在科学研究和各种活动中起继承、借鉴或参考作用的有价值的信息或知识。

情报学就是研究情报的产生、传递和利用规律的一门学科，现代情报学尤其关注利用现代化信息技术与手段，使情报交流过程、情报系统保持最佳效能状态，并帮助人们充分利用信息技术和手段提高情报产生、加工、储存、检索、交流、利用的效率。

当然，国内外情报界对情报和情报学的认识多种多样，尚无统一标准，认为情报学“是研究情报传递工作的理论、方法和原理的学科”，“是研究如何开发、利用信息资源的规律及其技术、方法的学科”或“是研究人类如何创造、利用和交流各种形式的情报的学科”等观点均有代表性。但总的来看，情报学可以说是以情报工作的理论、原则和技术、方法为研究对象的学科。

作为一门近半个世纪才发展起来的新兴学科，情报学的主要研究任务包括：①揭示情报和情报工作的本质与特性；②探索情报工作的产生和发展规律；③探讨情报工作的内容环节及其在情报工作中的地位、作用，它们之间的内在联系以及各个环节的原则和方法；④研究各个时期情报工作的方针、政策、标准化的内容，以及现代化理论和技术；⑤阐明情报工作的科学管理理论与方法，等等。因此，情报学研究具有交叉学科或跨学科特点，是一门综合性的横断学科。中国情报界也有人把“情报学”和“情报科学”看作两个不同的概念，认为情报学是研究情报工作的原理、工作方式与技术方法，情报的组织、管理、服务、评价及情报事业发展历史的学科。而情报科学则是综合研究由语言学、文字学、语文学、符号学、目录学、文献学、档案学、图书馆学、情报学等多学科知识的科学。

在中国,科技情报工作虽然早在20世纪50年代就与世界同时起步,但情报学作为一个独立的专业研究领域,却是1978年后随着科技情报专业教育的发展才得以确立的。因此,情报专业教育可以说是情报学的引擎。

为持续推动情报专业教育发展,浙江大学和上海交通大学的有关专业教师在吸收前辈老师和同辈学者的众多研究成果基础上,合作编写了这本《情报学基础教程》。

为避免繁杂的概念讨论,本书以三个简明扼要的科学化概念为基础:把信息作为物理实在并定义为信息密度的积分,将情报定义为有价值的信息(有效信息),再将知识定义为有效信息的积分。

在采用新的概念系统的同时,本书在体系组织方面的构造思路如下:

学习情报学,首先应当建立科学思维和量化观念,故将情报学经验规律作为第1章内容,这是文献计量学的一个子集,主要陈述布拉德福定律、洛特卡定律、齐夫定律、文献增长律、文献老化律的基本内容,并简要介绍其机理研究。

作为一门学科,应当有自己独立的基础理论体系,故第2章在扼要介绍国内外代表性情报学基础理论上,尝试用“一个轴心、三条原则、一组方程”构造情报学的基本理论轮廓和体系框架,使学生或读者能对情报学的总体结构有一个简明的整体把握。

情报学研究需要采用各种方法,除成熟的科学通用方法如逻辑分析法、数学分析法等外,也有本学科特有的引文分析法等独特方法,这构成本书第3章的核心内容。

情报源是情报工作和情报学研究的物质基础,故作为本书第4章。

围绕分类编目、主题标引、文摘索引等情报工作形成的情报管理不仅是情报工作的中心内容,而且是学习和研究情报学必不可少的基础,故本书第5章进行系统陈述。

情报检索不仅是情报学的核心内容之一,而且是当代本科生、研究生、科研人员应当掌握的基本技能,故本书第6章对其理论基础和主要操作技能进行介绍。

情报分析也是情报学的核心内容之一,是最能体现情报学“研究”价值的分支,具有开拓潜力和发展前景,值得重视,这就形成本书第7章的实体内容。

伴随情报学进化,情报技术也日新月异地发展,尤其是在计算机技术、网络技术、通信技术等的组合推进下,情报技术正发生翻天覆地的变化,本书第8章概要介绍现代主流情报技术。

此外,情报系统与情报服务、情报学与其他学科交叉形成的研究领域等内容也是情报学的重要组成部分,这就构成本书后两章。

本书具有一定逻辑结构层次的编排希望能为初学情报学的学生或读者提供一个比较明晰的情报学概貌,适合用作研究生、本科生、专升本等类型学生的“情报学基础”或“情报学概论”类课程教材,特别适合具有其他学科背景的学生尤其是具有理、工、农、医专业背景的学生学习情报学使用,也可供图书情报工作人员参考。全书以业界公认的知识内容为主体,也适当增加了独立探索研究成果。本书的特色是带有较强的科学实证思维方式,这也正是浙江大学和上海交通大学的学术传统,希望能有一定启发价值。

目录

第二版前言
第一版前言

第1章

情报学经验规律	1
1.1 布拉德福定律	1
1.2 洛特卡定律	4
1.3 齐夫定律	7
1.4 文献增长律	9
1.5 文献老化律	10
1.6 情报学经验规律的统一模型和机理研究	11
习题和思考题	15

第2章

情报学基础理论	16
2.1 国外代表性情报学基础理论	16
2.2 国内代表性情报学基础理论	21
2.3 DIKW 概念链及其量化	22
2.4 情报学学科架构	28
习题和思考题	31

第3章

情报学方法论	32
3.1 研究方法概述	33
3.2 通用方法概要	35
3.3 特色方法导引	46
习题和思考题	52

第4章

情报源	53
4.1 情报源导引	53
4.2 文献情报源	54
4.3 非文献情报源	79
4.4 开放存取情报源	81
习题和思考题	84

第5章

情报组织	85
5.1 分类组织法	85
5.2 主题组织法	93
5.3 索引文摘法	101
5.4 元数据	105
5.5 本体论	111
5.6 语义网	114
习题和思考题	118

第6章

情报检索	119
6.1 检索理论模型	119
6.2 检索评价参数	120
6.3 检索系统设计要点	121
6.4 检索策略设计要点	128
6.5 主要检索系统类型及评价	129

6.6 情报检索系统发展特点和趋势	135
习题和思考题	142
第7章	
情报分析	143
7.1 情报分析的类型	144
7.2 情报分析的操作原则	145
7.3 情报分析的流程和步骤	146
7.4 计算机辅助情报分析	151
7.5 跟踪型情报分析	151
7.6 评价型情报分析	154
7.7 预测型情报分析	162
习题和思考题	168
第8章	
情报技术	169
8.1 数据库及数据仓库技术	169
8.2 数据挖掘及文本挖掘技术	179
8.3 情报支持技术	189
习题和思考题	194
第9章	
情报服务	195
9.1 情报服务的模式	196
9.2 情报服务内容	201
9.3 情报服务的用户研究	211
9.4 情报服务评估	212
习题和思考题	216
第10章	
情报教育	217
10.1 情报学专业教育的兴起	218

10.2	中国的情报学专业教育·····	221
10.3	国外情报学教育的发展情况·····	229
10.4	各国情报学高等教育的课程体系·····	237
10.5	我国情报学继续教育和信息素养教育·····	242
	习题和思考题·····	249
第 11 章		
	当代情报学发展·····	250
11.1	信息资源管理·····	250
11.2	知识管理和竞争情报·····	255
11.3	科学计量学·····	264
11.4	信息经济学·····	267
11.5	信息伦理学·····	271
11.6	当代情报学研究前沿及未来研究趋势概述·····	275
	习题和思考题·····	277
	参考文献·····	278
	附录 学术规范概要·····	288
	第二版后记·····	293
	第一版后记·····	294



情报学经验规律

本章提要：

本章概述布拉德福定律、洛特卡定律、齐夫定律三大定态规律和文献增长律、文献老化律两类动态规律，并引介其统一机理。

主要知识点：

1. 布拉德福定律、洛特卡定律、齐夫定律三大定态经验定律；
2. 文献增长律、文献老化律两类动态经验规律；
3. 情报学经验规律的统一机理。

经验规律是学科赖以建立的基石，也是定量化研究的起点。情报学在经验规律方面已有一定积累，其基础主要由下列信息计量规律（Rousseau, 2010a）奠定。考虑到有专门的文献计量学或信息计量学教程（Egghe and Rousseau, 1990；王崇德，1990；丁学东，1993；邱均平，2001），本章在此对诸定律仅作简要介绍。

1.1 布拉德福定律

布拉德福定律是描述专业论文在期刊中分布情况的经验规律，由英国著名文献学家布拉德福（S. C. Bradford）于1934年提出。

1.1.1 基础数据和规律表述

布拉德福毕业于英国伦敦大学，所学专业是化学，1922年获科学博士学位。但他热爱图书馆工作和图书馆事业，1925~1937年曾长期担任英国南肯辛顿科学图书馆馆长。在工作中他发现一个学科的专业论文常分散发表在各种期刊上，于是他以应用地球物理学和润滑专业为实例，研究了期刊论文分布的情况，提出了一个描述文献分散的经验定律。布拉德福采集的1929~1932年应用地球物理学原始数据如表1-1所示。

表 1-1 布拉德福原始数据

期刊数	相关论文数	累积期刊数	累积论文数
1	93	1	93
1	86	2	176
1	56	3	235
1	48	4	283
1	46	5	329
1	35	6	354
1	28	7	392
1	20	8	412
1	17	9	429
4	16	13	493
1	15	14	508
5	14	19	578
1	12	20	590
2	11	22	612
5	10	27	662
3	9	30	689
8	8	38	753
7	7	45	802
11	6	56	868
12	5	68	928
17	4	85	996
23	3	108	1065
49	2	157	1163
169	1	326	1332

资料来源: Bradford, 1950; Garfield, 1980

如果将以上数据分成 3 个区域:

分区	期刊数	相关论文数	每刊年载相关论文数
c	$9 (n_c)$	429	>4
1	$59 (n_1)$	499	1-4
2	$258 (n_2)$	404	1

则近似有

$$n_c : n_1 : n_2 = 1 : a : a^2, a \approx 5$$

显然这只是一条近似规律。

用文字表述就是: 如果将期刊按其刊载某专业论文数量的多寡以递减顺序排列, 则可分出一个核心区和相继的几个区域, 当每区刊载的论文量相等时, 核心期刊数 n_c 和

外围一区期刊数 n_1 、外围二区期刊数 n_2 成 $n_c : n_1 : n_2 = 1 : a : a^2$ 关系。其中 a 称为布拉德福常数。

若以累积期刊数 n 或其对数为横坐标、累积相关论文数 $R(n)$ 为纵坐标，一般可作出如图 1-1 所示的分布曲线。

后来布鲁克斯 (B. C. Brookes) 用公式将该曲线对应的布拉德福定律表述为

$$R(n) = \begin{cases} \alpha n^\beta, & (1 \leq n \leq n_c) \\ k \lg(n/s), & (n_c \leq n \leq N) \end{cases} \quad (1-1)$$

其中, $R(n)$ 为相关论文累积数; N 是期刊总数, n 为期刊等级排序后的序号; α 是 $n=1$ 对应的 $R(n)$; β, k, s 为参数。曲线右上端的弯折称为格鲁斯 (Q. V. Groos) 下降。

1.1.2 改进和扩展

布拉德福 1934 年提出上述经验规律后, 并未引起人们注意。1948 年, 其专著《文献学》(Documentation) 出版, 在第 9 章中全文收录了 1934 年的发现, 才引起另一位英国文献学家维克利 (B. C. Vickery) 的重视。维克利对之进行研究后, 将外围区扩展到多个区域, 并首次用“布拉德福分散定律”命名这一经验规律 (Vickery BC and Vickery A, 1987)。

把期刊按照刊载某学科相关论文的数量排列, 可得到专门登载该学科论文的核心区期刊和若干外围区期刊, 当各区所含该学科相关论文相等时, 各区期刊数成

$$n_c : n_1 : n_2 : \dots = 1 : a : a^2 : \dots (a > 1) \quad (1-2)$$

其中, a 是布拉德福常数。并有一个改进称为维克利推论:

$$n_c : (n_c + n_1) : (n_c + n_1 + n_2) : \dots = 1 : b : b^2 : \dots (b > 1) \quad (1-3)$$

其中, b 叫做维克利系数。

1977 年, 苏联情报学家斯莫里科夫 (И. А. СМОЛЬКОВ) 对布鲁克斯的公式描述进行研究后, 提出用一个统一的方程来表述布拉德福定律:

$$R(n) = K \lg(q_1 n + q_2 e^{-\beta n}) \quad (1-4)$$

其中, K, β, q_1, q_2 均为参数。

1.1.3 应用提示

布拉德福定律的主要用途是确定核心期刊, 以指导期刊订购和期刊利用, 并由此扩展到核心馆藏维护、核心检索工具选择等。有人认为布拉德福定律可以看作是社会科学中普遍存在的“二八律”的一种表现, 即 20% 的核心期刊上刊载了 80% 的重要论文。

此外, 布拉德福定律也应用于考察专著的分布等。上海交通大学王国龙等通过对

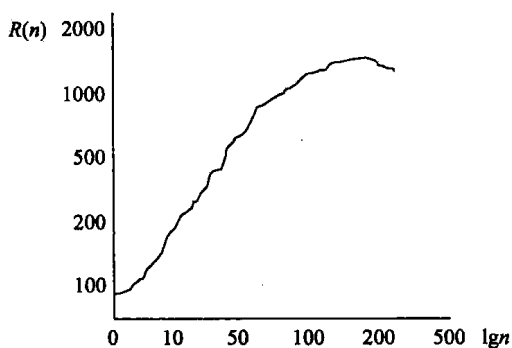


图 1-1 布拉德福分布曲线示意图

SCI 和 SSCI 引用期刊的研究,发现影响因子 (IF: Impact Factor) 的分布也符合布拉德福定律^①。

1.2 洛特卡定律

洛特卡定律是描述作者与其发表论文数量之间关系的经验规律,由美国统计学家洛特卡 (A. J. Lotka) 于 1926 年提出 (Lotka, 1926)。

1.2.1 基础数据和规律表述

洛特卡出生在波兰,曾留学法国、德国、英国,后在美国工作,出任过美国统计学会理事长。他的原创性研究是对物理学和化学两大学科中科学家们发表论文的情况进行的统计分析。对于物理学,他使用德国奥尔巴赫 (Aürbach) 《物理学史一览表》(Geschichtstafelen der Physik) 的人名索引,对其中作者(总计 1325 位)进行统计分析,依次列出发表 1 篇、2 篇和多篇论文的作者数;对于化学,他采用《化学文摘》(Chemical Abstracts, 1907—1917) 十年累积索引中的姓氏对以 A、B 开头的 6891 位作者进行了统计,分别列出发表 1 篇、2 篇、3 篇直至 346 篇论文的作者数。结果发现物理学 1325 位作者中,发表 1 篇论文的为 784 人,占 59.2%;化学 A 字母姓氏作者 1543 人,发表 1 篇论文的为 890 人,占 57.7%;B 字母姓氏作者 5348 人,写 1 篇论文的为 3101 人,占 57.98%。于是他提出作者的百分比分布(科学生产率的频率分布)应符合如下公式:

$$f(x) = \frac{c}{x^a} \quad (1-5)$$

其中, $f(x)$ 是发表 x 篇论文的作者占作者总数的百分比(作者频率),常数 $a > 1$ (经验研究表明 $1.8 < a < 3.8$, 估计浮动范围 $1.5 < a < 4$)。上式取对数,有

$$\log f(x) = -a \log x + \log c \quad (1-6)$$

故以 $\log f(x)$ 为纵坐标、 $\log x$ 为横坐标作图应是一条斜率为 $-a$ 、截距为 $\log c$ 的直线。

洛特卡的原始数据中, Aürbach 数据部分见表 1-2 (其中计算百分比是根据洛特卡平方反比律算得)。

表 1-2 洛特卡原始数据片段

论文数	统计数据			作图数据	
	撰写人数/人	实际比例/%	计算比例/%	数据点个数	斜率
1	784	59.17	60.79	13	-2.0411
2	204	15.40	15.20	14	-2.0533
3	127	9.58	6.75	15	-2.0182
4	50	3.77	3.80	16	-2.0255

^① 王国龙,李佩. SCI 源期刊与 JCR. 上海交通大学学报, 2003 (增刊): 287-290; 王国龙. SSCI 源期刊概述. 中国索引, 2004, 2 (2): 20-25

续表

论文数	统计数据			作图数据	
	撰写人数/人	实际比例/%	计算比例/%	数据点个数	斜率
5	33	2.49	2.43	17	-2.0210
6	28	2.11	1.69	18	-2.0953
7	19	1.43	1.24	19	-2.1385
8	19	1.43	0.95	20	-2.0786
9	6	0.45	0.75	21	-2.0726
10	7	0.53	0.61	22	-1.9887
11	6	0.45	0.50	23	-1.9989
12	7	0.53	0.42	24	-1.9946
13	4	0.30	0.36	25	-1.9754
14	4	0.30	0.31	26	-1.9490
15	5	0.38	0.27	27	-1.8485
⋮	⋮	⋮	⋮	⋮	⋮

资料来源：邱均平，1988

基于这些数据绘制的 $\log f(x) \sim \log x$ 直线如图 1-2 所示。

既然大部分数据点（尤其是前 13~25 个点）获得的斜率都接近 -2，故洛特卡定律近似平方反比律：

$$f(x) = \frac{c}{x^2} \quad (1-7)$$

其中 c 是常数。

$$\text{由于 } 1 = \sum_{n=1}^{\infty} f(n) = \sum_{n=1}^{\infty} \frac{c}{n^2} = c \frac{\pi^2}{6}, \text{ 故 } c = \frac{6}{\pi^2} \approx$$

60.79%， $f(1) = \frac{c}{1^2} = c$ ，进而有

$$f(2) = \frac{c}{2^2} = \frac{f(1)}{2^2} \quad (1-8)$$

$$\vdots$$

$$f(n) = \frac{c}{n^2} = \frac{f(1)}{n^2} \quad (1-9)$$

也就是说，发表 1 篇论文的作者约占作者总数的 60.79%，发表 2 篇论文的作者数是发表 1 篇论文作者数量的 $1/4$ ……发表 n 篇论文的作者数是发表 1 篇论文作者数量的 $1/n^2$ 。

洛特卡定律描述了作者人数与其发表论文章数之间的关系，首次揭示了作者与发表论文章数之间存在的规律。后经研究，发现物理学等学科领域的作者与论文章数之间的关系基本符合平方反比律，而生物、工程、计算机等领域则不符合平方反比律。一般来

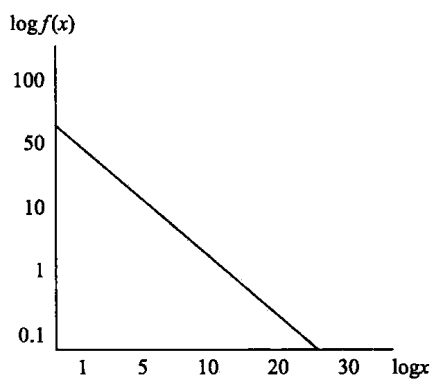


图 1-2 洛特卡分布曲线示意图

说,人文科学、社会科学中, a 值将变大; 规模较大、科研合作程度较高的学科中, a 值会变小。

1.2.2 改进与扩展

虽然洛特卡定律早在 1926 年就提出, 但直到 1949 年才被称为“洛特卡定律”。后来, 不少学者深入探讨了洛特卡定律的形成机理和改善表达, 著名科学家普赖斯 (D. Price) 就在洛特卡定律基础上进一步研究了科学家人数和科学论文数量之间的关联, 在《小科学, 大科学》一书中提出了“普赖斯定律”: 科学家总数是杰出科学家人数的平方 (或者说杰出科学家人数是科学家总数的平方根)。普赖斯定律的另一定量表达是: 杰出科学家发表了全部科学论文的 50%。

设 n_{\max} 为杰出科学家中最高产作者发表论文数, m 为杰出科学家中最低产作者发表论文数, R 为杰出科学家人数与全体科学家总数之比, 普赖斯还推导出以下两式:

$$m = 0.749\sqrt{n_{\max}} \quad (1-10)$$

$$R = 0.812/\sqrt{n_{\max}} \quad (1-11)$$

也就是说, 杰出科学家中最低产作者发表论文数是最高产作者发表论文数平方根的 0.749 倍; 杰出科学家人数与全体科学家总数之比是杰出科学家中最高产作者发表论文数平方根倒数的 0.812 倍。

比利时著名信息计量学家 L. Egghe 更进一步在洛特卡定律基础上发展出洛特卡信息计量学 (Egghe, 2005), 该学说具有较高的理论价值。

1.2.3 应用提示

洛特卡定律主要用于研究“科学生产率”, 可用于预测发表不同篇数论文的作者数量和特定学科的论文总量, 或根据作者数量估计科学论文数量等。

为合理评价洛特卡定律的适用性, 美国情报学家科依尔 (R. C. Coile) 于 1977 年提出用 K-S (Kolmogorov-Smirnov) 检验法对其进行鉴定, 其步骤为

(1) 设 A 为统计的作者总数, 计算 K-S 值: $K-S=1.63/\sqrt{A}$ 。

(2) 计算最大偏差值 D : $D=\max|F_0(x)-S_n(x)|$; 其中 $F_0(x)$ 为累积作者频率理论值, $S_n(x)$ 为累积作者频率观察值。

(3) 比较 D 与 K-S: 若 $D < K-S$, 则抽样分布符合洛特卡定律; 若 $D > K-S$, 则抽样分布不符合洛特卡定律。

科依尔用该方法检测了洛特卡《物理学史一览表》和《化学文摘》的原始数据, 发现《物理学史一览表》的数据符合洛特卡定律, 而《化学文摘》的数据则不完全符合洛特卡定律。科依尔还用该方法对美国国会图书馆的 MARC 作者统计数据 and 伊利诺依大学图书馆学院卡片目录作者统计数据进行了检验, 结果发现美国国会图书馆的 MARC 作者统计数据不符合洛特卡定律, 而伊利诺依大学图书馆学院卡片目录作者统计数据则完全符合洛特卡定律。K-S 检验法从此成为检验洛特卡定律适用性的有效方法。

国内也有不少洛特卡定律研究成果, 如福建农林大学张贤澳围绕洛特卡定律发表过