

第三版前言

《情报学基础教程》作为情报学教学单位的教学用书和教学参考书，先后于 2006 年出版第一版、2012 年修订第二版，并得到使用单位的欢迎和肯定，承蒙学界关注和出版社厚爱，现决定推出第三版。

这次修订仍延续理论联系实际、贯通学科基础与学科前沿的总体原则，适当强化教学逻辑。

各章修订思路如下。

第 1 章阐述情报学经验规律，旨在确立情报学的科学基础，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更新统一机理。

第 2 章阐述情报学基础理论，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，因而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概念的基础进行梳理，把情报学理论建立在量化 DIKW 概念链的基础上。

第 3 章阐述情报学方法论，根据研究型教学的需要重新改写，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，又为各章中微观层面的具体方法展开预留阐述空间。

第 4 章阐述情报信息源，作为情报工作和情报学研究的客观基础，修订版进行了改写和扩充。

从第 5 章到第 10 章，依次阐述情报组织、情报检索、情报分析、情报技术、用户行为与情报服务、情报教育，它们构成情报学各实体分支的主要内容，也是情报学核心课程的教学内容，本次修订侧重简明扼要地阐明其知识内容和研究动向，并全新改写了第 9 章。

第 11 章到第 12 章阐述行业情报和当代情报学发展，其中在本次修订中另辟新章专述行业情报，旨在为研究型情报学教学提供更多素材，并大幅扩充当代情报学发展内容，以展示情报学发展前景，同时，强化对大数据情报学的适当关注，也构成本次修订的另一特点。

修订和增改各章在阐明情报学基础知识的同时，力图展示情报学研究与发展。尽管各章内容相对独立，但保持了理论内核严谨一致，并通过探索引导面向前沿研究。作为情报学基础或情报学概论的教材，也期望能为后续专门课程的学习奠定坚实基础。

希望第三版继续保持把公认知识内容与独立专门研究相结合的特色，体现情报学理论知识与情报工作实践创新的有机交融，为普及和传播情报学知识贡献力量。

目 录

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..... 11
1.6	情报学经验规律的统一模型和机理研究..... 13
	习题和思考题..... 18
	参考文献..... 18
第 2 章	
	情报学基础理论..... 21
2.1	国外代表性情报学基础理论..... 21
2.2	国内代表性情报学基础理论..... 26
2.3	DIKW 概念链及其量化..... 27
2.4	情报学基本原理和体系结构..... 33
	习题和思考题..... 39
	参考文献..... 39
第 3 章	
	情报学方法论..... 42
3.1	研究方法总论：研究程式与学术规范..... 43
3.2	通用方法概要..... 51
3.3	特色方法导引..... 62
	习题和思考题..... 68
	参考文献..... 68

第4章

情报信息源	70
4.1 情报信息源概述	70
4.2 文献情报信息源	71
4.3 非文献情报信息源	98
4.4 开放存取信息资源	101
习题和思考题	105
参考文献	105

第5章

情报组织	106
5.1 分类组织法	106
5.2 主题组织法	114
5.3 索引文摘法	123
5.4 元数据	128
5.5 关联数据	134
5.6 本体和语义网	140
习题和思考题	145
参考文献	145

第6章

情报检索	147
6.1 情报检索基本原理	147
6.2 经典检索模型	149
6.3 检索系统设计要点	150
6.4 检索策略设计要点	156
6.5 情报检索评价	158
6.6 主要检索系统类型及评价	160
6.7 情报检索系统发展特点和趋势	165
习题和思考题	174
参考文献	174

第7章

情报分析	176
7.1 情报分析的操作原则	178
7.2 情报分析的流程和步骤	178
7.3 特定领域情报分析	184

7.4 情报分析工作类型	192
7.5 大数据时代情报分析的发展	204
习题和思考题	205
参考文献	205

第 8 章

情报技术	209
8.1 数据库及数据仓库技术	209
8.2 数据挖掘及文本挖掘技术	219
8.3 大数据及其处理技术	228
8.4 情报支持技术	233
习题和思考题	236
参考文献	236

第 9 章

用户行为与情报服务	237
9.1 以用户为中心的情报服务	237
9.2 用户信息需求与行为模型	239
9.3 现代情报服务内容体系	250
9.4 当代情报服务发展趋势	259
习题和思考题	262
参考文献	263

第 10 章

情报教育	268
10.1 情报学专业教育的兴起	269
10.2 中国的情报学专业教育	272
10.3 国外情报学教育的发展情况	280
10.4 各国情报学高等教育的课程体系	288
10.5 情报素养与情报素养教育	294
习题和思考题	297
参考文献	297

第 11 章

行业情报	299
11.1 行业情报概述	299
11.2 现代科技情报	300

11.3	商业情报与竞争情报	309
11.4	数据环境下的公安情报	317
11.5	行业情报前沿研究动向	322
	习题和思考题	325
	参考文献	325
第 12 章	当代情报学发展	328
12.1	知识管理和竞争情报	328
12.2	科学计量学与学术评价	338
12.3	信息可视化与科学知识图谱	345
12.4	信息经济学与政策分析	352
12.5	网络信息管理与商业数据分析	355
	习题和思考题	359
	参考文献	359
第三版后记		363

第1章

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，并引介这些规律的统一机理。

主要知识点

1. 布拉德福定律、洛特卡定律、齐夫定律三大定态规律。
2. 文献增长律、文献老化律两类动态规律。
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，2010）奠定。考虑到已有专门的文献计量学或信息计量学教程（王崇德，1990；丁学东，1993；邱均平，2001），本基础教材仅略为引导，以利后续专门学习。

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的定律，由英国著名文献学家布拉德福（S.C.Bradford，1878—1948）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	179
1	56	3	235
1	48	4	283
1	46	5	329
1	35	6	364
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	1 065
49	2	157	1 163
169	1	326	1 332

资料来源：Bradford（1950）；Garfield（1980）

如果将以上数据分成 3 个区域，则得出：

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	$1\sim4$
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领域，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，一般可做

出如图 1-1 所示的分布曲线。

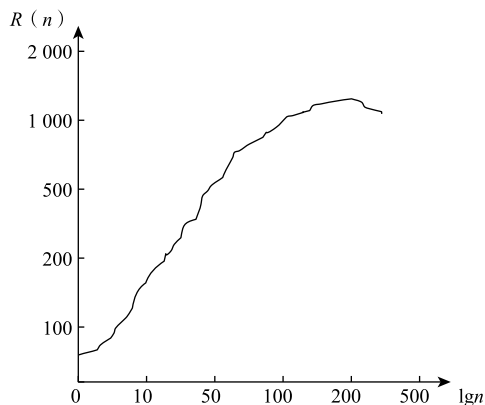


图 1-1 布拉德福分布曲线示意

后来布鲁克斯 (B. C. Brookes, 1990—1991) 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 为期刊总数; n 为期刊等级排序后的序号; α 为 $n=1$ 对应的 $R(n)$; β, k, s 为参数。曲线右上端的弯折称为格鲁斯下降 (Gross drop)。

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, 在第九章中全文收录了 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。维克利对其进行研究后, 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:

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, 各区期刊数可表述为

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

其中, a 为布拉德福常数。将其进行改进 (现称维克利推论) 则有

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

其中, b 为维克利系数。

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 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:20%的核心期刊上刊载了80%的相关论文。

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《科学引文索引》(*Science Citation Index*, SCI)和《社会科学引文索引》(*Social Science Citation Index*, SSCI)引用期刊的研究,发现影响因子(*impact factor*, IF)的分布也符合布拉德福定律(王国龙和李佩,2003;王国龙,2004)。

1.2 洛特卡定律

洛特卡(A. J. Lotka, 1880—1949)定律是描述作者与论文数量之间关系的经验规律,由美国统计学家洛特卡于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发表论文的情况进行的统计分析(statistical)。对于物理学,他使用德国奥尔巴赫(Aürbach)《物理学史一览表》(*Geschichtstafeln der Physik*, 1910年)的人名索引(index),对其中全部1325位作者进行统计分析,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年)11年累积索引中的姓氏以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, $f(x)$ 为发表 x 篇论文的作者占作者总数的频率(百分比);常数 $a > 1$ (经验研究表明 $1.8 < a < 3.8$,估计浮动范围 $1.5 < a < 4$)。上式取对数为

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

故以 $\log f(x)$ 为纵坐标、 $\log x$ 为横坐标作图应是一斜率为 $-a$ 、截距为 $\log c$ 的直线(本书 $\log$ 表示计算时取常用对数,表示分析时取自然对数)。

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	实际百分比	计算百分比	数据点个数	斜率
1	784	59.17%	60.79%	13	-2.041 1
2	204	15.40%	15.20%	14	-2.053 3
3	127	9.58%	6.75%	15	-2.018 2
4	50	3.77%	3.80%	16	-2.025 5
5	33	2.49%	2.43%	17	-2.021 0
6	28	2.11%	1.69%	18	-2.095 3
7	19	1.43%	1.24%	19	-2.138 5
8	19	1.43%	0.95%	20	-2.078 6
9	6	0.45%	0.75%	21	-2.072 6
10	7	0.53%	0.61%	22	-1.988 7
11	6	0.45%	0.50%	23	-1.998 9
12	7	0.53%	0.42%	24	-1.994 6
13	4	0.30%	0.36%	25	-1.975 4
14	4	0.30%	0.31%	26	-1.949 0
15	5	0.38%	0.27%	27	-1.848 5
...	...	...	...	...	...

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。

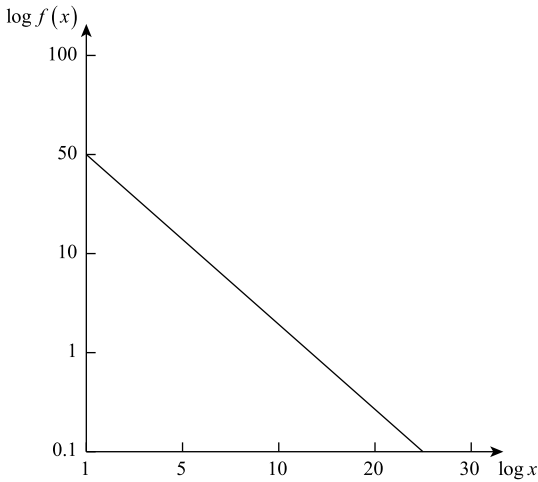


图 1-2 洛特卡分布曲线示意

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：

$$f(x)=\frac{c}{x^2}$$

(1-7)

其中, c 为常数。

由于 $1 = \sum_{n=1}^{\infty} f(n) = \sum_{n=1}^{\infty} \frac{c}{n^2} = c \frac{\pi^2}{6}$, 故 $c = \frac{6}{\pi^2} \approx 60.79\%$, $f(1) = \frac{c}{1^2} = c$, 进而有

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

.....

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

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量之间的关系, 首次揭示了作者与发表论文数量之间存在的定量规律。后经研究, 发现物理学等学科领域的作者与论文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关系。一般来说, 人文学科、社会学科中, 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 改进和扩展

虽然洛特卡定律早在 1926 年就提出, 但是直到 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艾格进一步把洛特卡定律作为一般幂律分布, 将其发展成为整个信息计量学的基础, 建立了洛特卡信息计量学 (Egghe, 2005), 具有较高理论价值。

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, 或根据特定领域作者数量估计该领域相关科学论文数量等。

为合理评价洛特卡定律的适用性, 美国情报学家科尔 (R. C. Coile) 于 1997 年提出

用 K-S (Kolmogorov-Smirnov) 检验法对其进行鉴定, 其步骤如下所述。

(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 : $D=\text{Max}|F_0(x)-S_n(x)|$; 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。科尔还用该方法对美国国会图书馆的 MARC (machine readable cataloging, 机读目录格式) 作者统计数据 and 伊利诺伊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, 而伊利诺伊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多篇论文 (刘婧, 2004), 并出版有专著《广义洛特卡定律——估计、推论及其在管理中的应用》, 可供参考。

1.3 齐夫定律

齐夫定律是描述文献中用词 (words) 与其出现频次之间关系的经验规律, 由齐夫 (G. K. Zipf) 于 1935 年提出, 是一条与语言学密切相关的文献学规律, 因而也成为计量语言学的基础定律之一。

1.3.1 基础数据和规律表述

1935 年, 哈佛大学语言学教授齐夫统计了乔伊斯 (J. Joyce) 的中篇小说《尤利西斯》(Ulysses) 中使用的 29 899 个词汇, 按使用频率 (词频) 排列成序, 并对应编定词级——使用频率最高的词汇的词级为 1, 以此类推, 则词级和词频的乘积接近于常数。他的统计数据如表 1-3 所示。

表 1-3 齐夫定律原始数据

词级 r	词频 F	词级 $r \times$ 词频 F
10	2 653	26 530
20	1 311	26 220
30	926	27 780
40	717	28 680
50	556	27 800
100	256	25 600
200	133	26 600

续表

词级 r	词频 F	词级 $r \times$ 词频 F
300	84	25 200
400	62	24 800
500	50	25 000
1 000	26	26 000
2 000	12	24 000
3 000	8	24 000
4 000	6	24 000
5 000	5	25 000
10 000	2	20 000
20 000	1	20 000
29 899	1	29 899

资料来源: Zipf (1972); Garfield (1980)

表 1-3 中的词频 F 是绝对频数 (频次), 若采用相对频数 (频率) $f=F/N$, N 为文章所含词汇总量, 则以 $\lg f \sim \lg r$ 作图为一直线, 如图 1-3 所示。

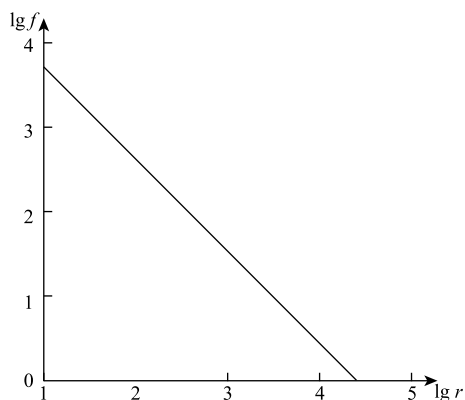


图 1-3 齐夫分布示意图

齐夫在 1949 年出版的《人类行为与最省力法则》(*Human Behavior and the Principle of Least Effort*) 专著中, 系统地将他的发现整理为一条规律。

设有一篇含有 N 个词的文献 ($N \geq 5\,000$), 用自然数 1, 2, 3, ... 给文献中的词编级, 出现频次最高的为 1, 其次为 2, 直到 r ($r < N$), 则高频词的频率 f 与词级 r 的乘积是一个常数。

$$fr = c \quad (1-12)$$

其中, $1 < c = C/N < \infty$, $Nc = C$ 亦为常数, 这就是齐夫定律。齐夫本人用最省力法则作为该定律的机理——使用频率高的词实际价值不大, 只是传递起来省力, 因此词频与词级的乘积基本稳定在一个常数水平。

1.3.2 改进和扩展

1936年,美国语言学家朱斯(M. Joos)提出含两个参数的齐夫公式。

$$fr^b = c \quad (1-13)$$

这是双参数词频分布律,其中 $b>0$, $c>0$ 。当 $b=1$ 时,等价于齐夫单参数公式。

1952年,美籍法裔数学家曼德布罗(B. Mandelbrot)运用信息论原理和概率论方法研究了词频分布规律,结合其独创的分形学说重新解释了齐夫定律,提出了三参数词频分布律。

$$f(r+a)^b = c \quad (1-14)$$

其中, $0 \leq a < 1$, $b>0$, $c>0$ 。该公式当 $a=0$ 时,就是朱斯双参数公式;当 $a=0$, $b=1$ 时,就是齐夫单参数公式,因而三参数公式具有更广的普适性,成为排序频次分布的一般形式。

但以上改善均针对高频词,低频词的分布规律是由英国学者布斯(A. D. Booth)首先导出的。

设出现1次的词数量为 I_1 ,出现 n 次的词数量为 I_n ,则低频词满足式(1-15)。

$$\frac{I_n}{I_1} = \frac{2}{n(n+1)} \quad (1-15)$$

式(1-15)也称为齐夫第二定律,式(1-12)相应称为齐夫第一定律。

1.3.3 应用提示

齐夫定律主要可用于指导文献标引、词表编制和情报检索等,尤其是在自动标引(automatic indexing)中具有应用价值,因为使用频率太高的词和使用频率太低的词都没有多少检索意义,故应当使用词频适中、词义强健的词作为标引词和检索词。

1.4 文献增长律

文献增长律是描述文献信息量随时间而不断增加的一组经验规律,主要有以下模型。

1.4.1 指数曲线

早在1949年,普赖斯就发现“十年一叠的《英国皇家学会哲学汇刊》靠墙竟堆成了一条完美的指数曲线”。1950年,他发表了第一篇关于“指数增长”的论文。随后,他对1665年《哲学汇刊》创刊以来世界科学杂志的增长数据、《化学文摘》《生物学文摘》《科学文摘》等摘录的期刊论文数量进行了统计,在1961年正式出版的《巴比伦以来的科学》中,系统地提出了科技文献按指数规律增长的经验规律。

设 t 时间的文献信息数量为 $Q=F(t)$,则

$$Q = F(t) = ae^{bt} = a \exp(bt) \quad (1-16)$$

其中, $a>0$ 为条件常数,代表 $t=0$ 时刻的 Q ; $b>0$ 为时间常数,代表持续增长率; e 为自然对数底(≈ 2.718); t 通常以年计算。

支持文献指数增长规律的部分数据参见表1-4。

表 1-4 世界期刊增长数据

时间	期刊种数
18 世纪 50 年代	10
19 世纪	100
19 世纪 50 年代	1 000
20 世纪	10 000
20 世纪 50 年代	50 000
21 世纪	100 000

资料来源：多种信息渠道汇总估计

参照表 1-4 数据，以公元年代作为横坐标时间 t ，以期刊数作为纵坐标文献量 $F(t)$ ，做成如图 1-4 所示曲线，即普赖斯曲线。

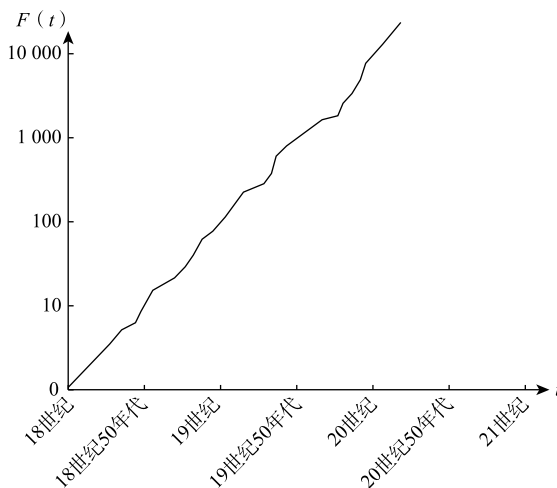


图 1-4 普赖斯曲线示意图

按照指数增长律，如果初始文献量 $a=1\ 000$ （件）， $b=10\%$ ，则：

10 年后文献量 $F(10)=1\ 000\exp(0.1\times 10)=2\ 718$ （件）

100 年后文献量 $F(100)=1\ 000\exp(0.1\times 100)=22\ 026\ 466$ （件）

因此，有“信息爆炸”之说。然而，实际上文献信息的增长是有限的，这表明指数增长律只是一条近似规律，故产生以下改进模型。

1.4.2 逻辑曲线

苏联科学学家纳里莫夫（В. Налимов）等在仔细研究指数增长律后，发现文献增长是分阶段进行的，在经过一段急剧增长过程后，会有一段缓慢增长，并有最终上限，也就是说指数曲线会变为逻辑曲线（一种生长曲线），其数学模型为

$$Q=F(t)=\frac{k}{1+a\mathrm{e}^{-bt}}=\frac{k}{1+a\exp(-bt)} \quad (1-17)$$

其中， k 为最大上界即 $t\rightarrow\infty$ 时的文献累积量。其形状如图 1-5 所示。