

科学文献的增长规律和老化规律及其新的一般模型

The Growth and Obsolescence of Scientific Literature and Their New General Models

林 辉

林 伟

(西北工业大学图书馆 西安 710072) (西北工业大学理学院 西安 710072)

摘 要 研究了科学文献增长规律和老化规律,利用微分方程法,分别提出了它们新的一般模型,还讨论了零增长问题。用一些包含著名模型的例子来说明新的模型具有良好的综合性。

关键词 科学文献 增长规律 老化规律 一般模型 零增长理论

中图分类号 G250

文献标识码 A

文章编号 1002-1965(2010)04-0022-04

众所周知,分析和研究科学文献变化的一些规律是人们关注的课题。多年来,经过人们的不断努力,已取得许多研究成果。在此基础上,本文研究了科学文献增长规律和老化规律,利用微分方程法,分别提出了它们新的一般模型,还讨论了零增长问题,并且用一些例子来说明新的模型具有良好的综合性。可以看到,例子中包含了一些著名的模型。

1 科学文献增长规律

1.1 科技文献的指数增长规律 美国著名学者普赖斯(Derek J de Solla Price)曾对科技文献增长规律进行了研究。他将1665~1900年每隔50年世界出版的科学杂志进行了统计分析。后来,他又统计分析《物理学文摘》等30种文献杂志中的期刊论文增长情况,提出了如下的观点:科技文献是按指数律增长的^[1]。如果用 $N(t)$ 表示时刻 t 的文献量,则指数增长规律可以表示为:

$$N(t) = a \exp[bt] \quad t \geq 0 \quad (1)$$

式中 a 表示统计初始时刻($t=0$)的文献量, b 表示增长率, $\exp[\cdot]$ 表示指数函数。或者,将指数增长规律表示为:

$$N(t) = a \exp[b(t-c)] \quad t \geq c \quad (2)$$

式中 a 表示统计时刻($t=c$)的文献量。显然,式(2)可以改写成

$$N(t) = A \exp[bt] \quad t \geq c \quad (3)$$

式中 $A = a \exp[-bc]$ 。容易看出,式(1)是下列方程的解

$$\frac{dN(t)}{dt} = bN(t) \quad (4)$$

1.2 科技文献增长规律的逻辑斯谛增长模型 大家知道,由于存在着环境限制,文献量不可能无限地增长。为了描述这种情况,前苏联学者费勒杜茨和纳利莫夫从逻辑斯谛增长模型出发,研究了文献增长规律^[2]。这时,需要将式(4)作些改变,即在该式的右端乘上一项 $(1 - N/K)$,得

$$\frac{dN}{dt} = bN(1 - N/K) \quad (5)$$

式中 K 是一个正的常数,称为容纳量(Carrying Capacity),它表示在该环境下所允许的最大文献量。设 $R = (1 - N/K)$,它代表着环境限制的效应。显然, $0 \leq R \leq 1$ 。式(5)的解为^[1]:

$$N(t) = K / (1 + a \exp[-bt]) \quad (6)$$

式中 $a = (K/N_0) - 1$, N_0 表示统计初始时刻文献的数量。它的曲线呈S型。显然,直线 $N(t) = K$ 是它的渐近线。

称式(5)或(6)为逻辑斯谛(Logistic)增长模型。应当说明,目前这一模型有不少的应用。

1.3 科技文献增长规律的舍-布模型 在指数增长模型和逻辑斯谛增长模型的基础上,前苏联学者舍斯托帕尔(V. M. Shestopal)和布勒曼(L. N. Burman)提出了科技文献增长的一般模型(General Model)如下^[2]:

$$\frac{dN(t)}{dt} = g(t) N(t) \quad (7)$$

式中 $g(t)$ 是一个时间函数,并且它是非增函数,

即 $g(t)$ 的导数小于或等于零。于是,式(7)的解为:

$$N(t) = A \exp\left[\int g(t) dt\right] \quad t \geq 0 \quad (8)$$

式中 A 是常数,它由初始条件来确定。

可以看出,指数增长模型、线性增长模型和逻辑斯谛增长模型和都是它的特殊情形。事实上,

a. 若 $g(t) = b$, 则式(7)变成式(5),这就是指数增长模型。

b. 若 $g(t) = b/(1 + bt)$, 其中 b 是正的常数,则将它代入式(8),可得

$$N(t) = A(1 + bt) \quad (9)$$

这是一条直线,它的斜率为 Ab 。这就是线性增长模型^[2]。

c. 若 $g(t) = b \{1 - 1/(1 + a \exp[-bt])\}$, 其中 a 和 b 都是正的常数,则经过简单的演算或查积分表^[4],可得

$$\int g(t) dt = -\ln(1 + a \exp[-bt]) \quad (10)$$

将它代入式(8),并令 $A = K, a = (K/N_0) - 1, K > N_0$, 于是,我们便得到式(6)。因此,逻辑斯谛增长模型是它的特殊情形^[5]。

1.4 科技文献增长规律的系统模型 2003年,王宏鑫从系统论观点出发,将舍-布模型中的函数 $g(t)$ 作了具体化假设,即设

$$g(t) = a_1 + a_2 N + a_3 N^2 + \dots + a_m N^{m-1} \quad (11)$$

式中 $N = N(t)$, m 是参数。于是,式(7)变成

$$\frac{dN(t)}{dt} = a_1 N + a_2 N^2 + a_3 N^3 + \dots + a_m N^m \quad (12)$$

这就是王宏鑫所提出的科技文献增长规律的系统模型^[6]。

1.5 科技文献增长规律的一般模型 由于存在着环境限制,文献量不可能无限地增长。因此,所要建立的模型应当符合这一客观要求。但是,可以发现,在上述的舍-布模型和系统模型中,无法保证每一个具体模型都符合这一要求。例如,它们都含有指数增长模型和线性增长模型^[6],而随时间不断地推移,这两者所描述的文献量却无限地增大。为了改变这种状况,下面我们提出一种新模型。

考察下面的方程

$$\frac{dN(t)}{dt} = f(t)(1 - N/K) \quad (13)$$

式中 K 是一个正的常数,它表示是容纳量,而 $f(t)$ 为非负函数,且

$$f(t) = a_1 N + a_2 N^2 + a_3 N^3 + \dots + a_m N^m \quad (14)$$

式中 a_1 是一个正的常数,其它系数都是常数。因此,从式(13)可以看出, $N(t)$ 为非减函数,且当 $N(t)$ 趋向 K 时, $N(t)$ 的变化率趋向零。于是,直线 $N(t) = K$ 是曲线 $N(t)$ 的渐近线。这说明,可以用它来描述文献量不可能无限地增长的情形。因此,我们称式(13)或它的解为科技文献增长规律的一般模型。

容易看出,若在式(14)中,令 $a_1 = b$ 并且其它系数都是零,则将它们代入式(13),便得到式(5)。于是,这说明了逻辑斯谛增长模型是它的特殊情形。

为了进一步说明上述的一般模型与舍-布模型和系统模型之间的关系,我们假设文献量可以无限地增长,即式(13)中的 K 可以无限地增大,因此 N/K 变成零,这时式(13)变成

$$\frac{dN(t)}{dt} = f(t)$$

此时

a. 令 $f(t) = g(t)N(t)$, 其中 $g(t)$ 是非负函数,于是此式变成式(7),即变成舍-布模型;

b. 令 $f(t) = g(t)N(t)$, 其中 $g(t)$ 是非负函数,且由式(12)表示,因此 $f(t) = a_1 N + a_2 N^2 + a_3 N^3 + \dots + a_m N^m$ (14)

式中 a_1 是一个正的常数,其它系数都是常数,于是方程变成式(12),即变成系统模型。

1.6 零增长理论 为研究社会经济发展问题,1972年,美国学者梅多斯(D. H. Meadows)等人发表《增长的极限》一书^[7],提出了零增长理论(Zero Growth Theory)。他们主要从自然条件、资源、环境等方面论证经济增长的“极限”。后来,美英学者应用零增长理论,研究了图书馆馆藏问题。他们提出了一种控制馆藏增长的理论,又称馆藏稳定状态理论,其中提出要建立有限规模的图书馆。也就是说,当图书馆发展到一定目标(馆藏量、功能等指标)之后,剔除文献的速度应等同于新购进文献的速度,使得馆藏文献资料实际增长速度为零,因而图书馆收藏的文献总量保持一种稳定状态^[8]。

虽然,国外学者对上述的馆藏量的零增长理论还有不同的看法,而我国一些研究者认为零增长理论对我国仍有借鉴的意义。如文献[8]认为:

a. 零增长理论可以作为我国图书馆文献信息资源建设与发展的指导思想之一。b. 零增长理论可保证我国图书馆文献资源建设稳定、和谐地发展。c. 零增长理论可以用来指导图书馆藏书的初选工作。d. 零增长理论也可以用来指导图书馆藏书的复选与剔除工作。e. 实践零增长理论,可促进文献信息资源共享。

现在有一个重要问题:对一个图书馆来说,在各

种客观条件不变的情况下,如何判定什么时候文献增长速度将趋向于零?为了说明如何求解这个问题,作为例子,不妨假设已知文献增长规律属于逻辑斯谛增长模型。这时,由于直线 $N(t) = K$ 是渐近线,因此,对于任何时刻 t 来说, $N(t)$ 的函数值不可能为 K 。于是,需要设定一个界限来表示馆藏文献总量将达到饱和状态。譬如说,假设这个界限为 $0.99K$,那么从式(6)可得

$$0.99K = K/(1 + a \exp[-bt]) \quad (15)$$

经过简单的运算,得

$$t = (1/b) \ln(99a) \quad (16)$$

这个结果表明,从时刻 $t = (1/b) \ln(99a)$ 起,文献增长速度接近于零。

2 科技文献老化规律

2.1 指数衰减规律 假设由于科技文献老化而引起被引用的文献量之衰减,因此从这个角度上来说,被引用的文献量衰减问题也是科技文献老化问题。

设最初时刻被引用文献量为 N_0 ,并且被引用文献量 $N(t)$ 的减少服从于指数衰减规律,即

$$N(t) = N_0 \exp[-bt] \quad t \geq 0 \quad (17)$$

式中 b 表示一个正的常数,并且 $0 < b < 1$ 。显然,它是下式的解

$$\frac{dN(t)}{dt} = -bN(t) \quad (18)$$

这就是英国学者布鲁克斯(Bertram C. Brookes)提出的指数衰减模型^[9-10]。它描述了科技文献被引用数量随时间推移而减少的规律。

为了绘图方便起见,通常将式(17)改写成

$$\ln N(t) = \ln N_0 - bt \quad t \geq 0 \quad (19)$$

可以看出,在对数坐标系中这是一条直线,它的斜率为 $-b$ 。

2.2 科技文献老化规律的靖-康模型 靖培栋和康仲远从逻辑斯谛增长模型出发,研究了科技文献老化规律的模型^[11]。他们将式(5)改写成

$$\frac{dN}{dt} = b(N - \frac{N^2}{K}) \quad (20)$$

或

$$\frac{dN}{dt} = b[N(t) - D(t)] \quad (21)$$

式中 $D(t) = N^2/K$, $N(t)$ 的最大值不超过 K ,而 K 是一个正的常数。在^[11]中称 $D(t)$ 为文献老化量。这时,由于 $N(t)$ 服从逻辑斯谛增长模型,因此,可以利用式(6)求得

$$D(t) = \frac{K}{(1 + a \exp[-bt])^2} \quad (22)$$

式中 $a = (K/N_0) - 1$, N_0 表示统计初始时刻($t = 0$)文献的数量。容易看出

$$D(0) = N_0, \lim_{t \rightarrow \infty} D(t) = K \quad (23)$$

由此可见,直线 $D(t) = K$ 是曲线 $D(t)$ 的渐近线。

2.3 科技文献老化规律新的一般模型 为了寻找科技文献老化规律的一般模型,我们用一个时间函数 $r(t)$ 来代替式(18)中的中常数 b 。于是,我们得到

$$\frac{dN(t)}{dt} = -r(t)N(t) \quad (24)$$

它的解为

$$N(t) = B \exp[-\int r(t) dt] \quad (25)$$

式中 B 为常数,它由初始条件来确定。我们称式(24)或(25)为科技文献老化规律的一般模型。另外,从式(24)还可求得下列关系式

$$r(t) = -\frac{d \ln N(t)}{dt} \quad (26)$$

需要注意,当研究被引用的文献量衰减问题时,式(24)中的 $r(t)$ 为非负函数,而当研究文献老化量增加问题时, $r(t)$ 为非正函数。

为了说明式(24)或(25)是科技文献老化规律的一般模型,我们进行下列的讨论:

a. 若 $r(t) = b$,则式(24)变成式(18)。这就是布鲁克斯的指数衰减模型。

b. 若 $r(t) = \frac{a \cdot \exp[-at] - ma \cdot \exp[-mat]}{\exp[-at] - \exp[-mat]}$,其中 a 和 m 都是正的常数, $m > 1$,则将它代入式(25),令 $B = C_0$,经过简单的演算或查积分表,可得

$$N(t) = C_0(\exp[-at] - \exp[-mat]) \quad (27)$$

如果 $N(t)$ 为引文频次, C_0 为传播幅度, a 为年代衰减率, m 为初始增量,那么这个式子就是罗马尼亚学者阿拉莫斯库(A. Avramescu)所提出的情报散播模型^[10,12]。

c. 若 $r(t) =$

$$\frac{a \cdot \exp[-t] + 2b \cdot \exp[-2t]}{1 - a \cdot \exp[-t] - b \cdot \exp[-2t]}$$

其中 a 和 b 都是正的常数, $a + b = 1$,则将它代入式(25),令 $B = 1$,经过简单的演算或查积分表,可得

$$N(t) = 1 - \exp[-t] - \exp[-2t] \quad (28)$$

如果 $N(t)$ 是以小数表示被引用文献之累计百分比, t 是被引用文献出版年龄(以十年为单位),那么这个式子就是美国学者伯顿(B. E. Burton)和开普勒(R. W. Keblor)所提出的公式^[2,10]。

d. 若 $r(t) =$

$$\frac{a \cdot \exp[-(t-0.1)] + 2b \cdot \exp[-(2t-0.2)]}{1 - a \cdot \exp[-(t-0.1)] - b \cdot \exp[-(2t-0.2)]}$$

其中 a 和 b 都是正的常数, $a + b = 1$, 则将它代入式(25), 令 $B = 1$, 经过简单的演算或查积分表可得

$$N(t) = 1 - \exp[-(t - 0.1)] - \exp[-(2t - 0.2)] \quad (29)$$

如果 $N(t)$ 是以小数表示被引用文献之累计百分比, t 是被引用文献出版年龄(以十年为单位), 那么这个式子就是前苏联学者莫德列夫的修正公式^[10]。

e. 若 $r(t) = -ab \exp[-bt]/(1 + a \exp[-bt])$, 其中 a 和 b 都是正的常数, 则经过简单的演算或查积分表, 可得

$$\int r(t) dt = \ln(1 + a \exp[-bt])$$

于是, 将它代入式(25), 可得

$$N(t) = \frac{B}{1 + a \cdot \exp[-bt]} \quad (30)$$

假设 $B = K$, $N(0) = N_0$, $a = (K/N_0) - 1$, $K > N_0$, 代入此式并与式(22)比较, 可得

$$N(t) = \sqrt{KD(t)} \quad (31)$$

这表明, 上述的科技文献老化的一般模型也可引出^[11]中的文献老化模型, 即式(22)。

3 结束语

在社会学中, 人们研究了人口的增长和减少问题, 这对社会发展是有意义的。同样地, 人们研究科学文献增长规律和老化规律, 对图书馆的发展是有益的。

本文研究了科学文献增长规律和老化规律, 在前人研究的基础上, 利用微分方程法, 分别提出了它们新

的一般模型。为了说明模型具有良好的综合性, 列举了一些例子。可以看到, 例子中包含了一些著名的模型。

参考文献

- [1] 吴慰慈主编. 图书馆学基础[M]. 北京: 高等教育出版社, 2004
- [2] 王崇德. 文献计量学引论[M]. 桂林: 广西师范大学出版社, 1994
- [3] V M Shestopal, L N Burman. The General Model for the Growth of Literature (in Russian)[J]. NTI, ser. 1, 1988(5): 23-28
- [4] Yi SU, Li-Feng HAN. A new Literature Growth Model: Variable Exponential Growth Law of Literature[J]. Scientometrics, 1998, 42(2): 295-265
- [5] 《实用积分表》编委会. 实用积分表[M]. 合肥: 中国科学技术大学出版社, 2006
- [6] 王宏鑫. 文献增长经典模型的系统论统一[J]. 情报杂志, 2003(5)
- [7] 德内拉·梅多斯(D. H. Meadows)等. 增长的极限[M]. 北京: 机械工业出版社, 2006
- [8] 徐恩元. 零增长理论研究[J]. 四川图书馆学, 2005, 146(4): 9-13
- [9] Bertram C Brookes. The Growth, Utility, and Obsolescence of Scientific Periodical Literature[J]. J. of Doc., 1970, 26(4): 283-294
- [10] 丁学东. 文献计量学基础[M]. 北京: 北京大学出版社, 1993
- [11] 靖培栋, 康仲远. 关于科技文献增长的数学模型[J]. 情报学报, 2000, 19(1): 90-96
- [12] A Aramescu. Actuality and Obsolescence of Scientific Literature[J]. J. ASIS, 1979, 30(5): 296-303

(责编: 白燕琼)

(上接第15页)

- Society for Information Science and Technology, 2007, 58(9): 1303-1309
- [3] 田大芳. 图书情报学期刊互引网络结构分析[J]. 情报杂志, 2009, 28(6): 48-51, 29
- [4] 邱均平, 赵为华. 期刊同被引的实证计量研究[J]. 情报科学, 2008, 26(10): 1447-1450
- [5] 刘则渊, 陈悦, 侯海燕. 科学知识图谱方法与应用[M]. 北京: 人民出版社, 2008, 8
- [6] 刘军. 社会网络分析导论[M]. 北京: 社会科学文献出版社, 2004
- [7] Evelien Otte, Ronald Rousseau. Social Network Analysis: A Powerful Strat-Egy. Also for the Information Sciences. Journal of Information Science, 2002, 28(6): 441-453
- [8] 徐媛媛, 朱庆华. 社会网络分析法在引文分析中的实证研究[J]. 情报理论与实践, 2008, 31(2): 184-188
- [9] 中国社会科学院文献信息中心文献计量学研究室. 中国人文社会科学核心期刊要览[M]. 北京: 社会科学文献出版社, 2008

- [10] 罗家德. 社会网分析讲义[M]. 北京: 社会科学文献出版社, 2005
- [11] 佟贺丰, 武夷山. 学术期刊应避免同质化[J]. 科技导报, 2009, 27(3): 107
- [12] 姜春林, 唐悦, 杜维滨等. CSCI 管理学来源期刊引文网络结构分析[J]. 科学与科学技术管理, 2009, 30(7): 54-54
- [13] 胡志刚, 李志红. 近十年我国科学学的学术群体与研究热点分析——基于9种科学学类期刊的科学计量学研究[J]. 科学与科学技术管理, 2009, 30(7): 13-18
- [14] 胡志刚, 侯海燕. 我国科学学期刊群的可视化分析[J]. 大连理工大学学报(社会科学版), 2009, 30(2): 120-123
- [15] 李树华. 科技期刊编辑的高境界: 洞见和齐物[J]. 编辑学报, 2009, 21(1): 89-90
- [16] 孙兆刚. 论我国科学学的演进图景[J]. 科学与科学技术管理, 2008, 29(1): 5-10
- [17] 刘则渊. 推进科学学的学科建设与前沿研究[J]. 科学与科学技术管理, 2005, 25(5): 18-20

(责编: 贺晓利)