

基于知识增长规律的知识存量测度

吴新文¹,熊永豪¹,赵 飒²,梁竞艳¹,马 晶³,黄晓玲³

(1.湖北省科技信息研究院,湖北 武汉 430071;2.中国地质大学,湖北 武汉 430074;3.湖北经济学院,湖北 武汉 430205)

摘 要:分析了两种经典模型——知识指数增长规律和逻辑增长规律,探讨了知识的老化现象及其影响因素,进而确定了知识老化率或折旧率的计算方法,并研究了基于科技文献的知识存量测度模型。

关键词:知识测度;指数增长;逻辑增长;知识折旧率;折旧因素

DOI:10.6049/kjbydc.2012070634

中图分类号:G302

文献标识码:A

文章编号:1001-7348(2013)16-0152-04

0 引言

知识增长具有一定规律性^[1]。关于知识增长规律的研究很早就引起了国内外学者的极大关注,这些学者对表征知识发展尤其是科学发展的各种计量指标,如科学刊物、科学论文、科学发明、科研机构、科技人力等进行了大量统计,取得了一系列知识增长规律方面的研究成果。借助知识增长规律,我们可以把握人类知识变化的基本趋势,进而衡量知识存量的基本状况,更好地认识和运用人类共同拥有的知识财富^[2]。

1 两种经典模型

1.1 知识指数增长规律

美国学者 F·赖德(F. Reyder)在探索知识增长规律时研究了美国主要大学图书馆藏书量的变化,发现藏书量平均每 16 年翻一番。随后,普赖斯(D. Price)综合分析了大量统计资料,他以横轴代表历史年代,以纵轴代表科技文献量,用一条光滑曲线连接这些坐标点,得出科技文献与时间的变化规律函数,即:

$$N = N_0 e^{bt} \quad (1)$$

可以看出,这是一个指数函数,其中 N_0 为初始文献量, b 为常数。

在科学学专著——《巴比伦以来的科学》中,普赖斯教授以科技文献(科技期刊和论文)衡量知识量,描述了知识发展的加速规律。1665 年,世界上最早的科学期刊——《哲学论坛》出版。在全世界范围内,科学期刊 1750 年接近 10 种,1800 年有 100 种,1850 年有

1 000 种,1900 年有 1 万种,到现在有 10 万多种。可以看出,科学期刊的数量从 1750 年起,大约每半个世纪增加 10 倍。

在另一本《小科学、大科学》名著中,普赖斯进一步概括了各种科学指标每过若干年总数翻一番的规律(见表 1)。

表 1 科学指标的翻番时间

| 成倍增长的周期(年) | 各种科学指标 |
|------------|-----------|
| 20 | 杰出的物理学家 |
| 20 | 重大的科学发现 |
| 20 | 化学元素的数量 |
| 20 | 仪器的精密度 |
| 20 | 每千人中大学生数目 |
| 15 | 科学期刊的种类 |
| 15 | 科学学会成员数目 |
| 15 | 化合物数量 |
| 15 | 各科科学文摘数量 |
| 15 | 文、理学士数量 |
| 10 | 小行星发现的数目 |
| 10 | 行列式理论文献数 |
| 10 | 非欧几何文献数 |
| 10 | 伦琴射线文献数 |
| 10 | 实验心理学文献数 |
| 10 | 交通速度 |
| 10 | 发电量 |
| 5 | 国际间电话通讯量 |
| 1.5 | 加速器能量 |

科技文献增长规律在很大程度上表征了人类科研活动的规律性,反映着人类科学知识增长规律及其继承性和积累性。但是,指数增长规律并不是在任何条件下都成立的。大量研究表明,所统计的学科范围越

收稿日期:2012-11-03

作者简介:吴新文(1964—),男,湖北汉川人,湖北省科技信息研究院研究员,研究方向为科技管理;熊永豪(1970—),男,湖北嘉鱼人,湖北省科技信息研究院高级会计师,研究方向为会计理论及实务;赵飒(1985—),女,辽宁沈阳人,中国地质大学硕士研究生,研究方向为教育管理。

广,符合指数规律的时间就越长,而在某一具体学科领域文献量的增长符合指数规律的时间就较短。下面对其作进一步分析。

若时间增量为 Δt ,则科技文献的增长量 ΔN 为:

$$\Delta N = N_0 e^{b(1+\Delta t)} - N_0 e^{bt} = N_0 e^{bt} (e^{b\Delta t} - 1)$$

若设 $\Delta t=1$,则有:

$$\Delta N = N_0 e^{bt} (e^b - 1)$$

若 $t \rightarrow 0$,则 $\Delta N \rightarrow \infty$ 。这表明,单位时间内文献增长量随着时间的推移将趋于无穷大。实际上在任何时候,其均不可能趋于无穷大。这就需要进一步思考科技文献增长的机理。

1.2 逻辑增长规律

指数增长曲线准确地揭示了知识增长规律,但仅限于初期阶段。B. 纳里莫夫和 Г. 弗莱杜茨等人研究用逻辑曲线(Logistic curve)来描述知识增长规律。逻辑曲线(见图 1)说明科学知识呈“S”型增长,在初期阶段,知识迅速增长;随后,其增长速度减慢,并逐渐趋向饱和极限。

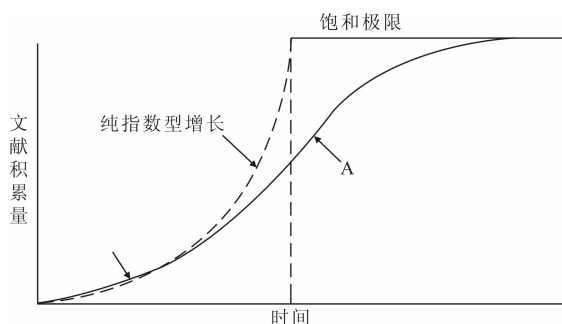


图 1 逻辑曲线

其数学表达式为:

$$N = N(t) = \frac{k}{1 + ae^{-kbt}} \quad (2)$$

其中, k 为当 $t \rightarrow \infty$ 时知识累积量的最大值, k, b, a 为大于零的常数, t 为时间变量, N 为时刻 t 科技知识的累积量。

式(2)对时间求导数,得到知识量增长速度:

$$\frac{dN}{dt} = \frac{k^2 abe^{-kbt}}{(1 + ae^{-kbt})^2} = kbN - bN^2 = bN(k - N) \quad (3)$$

通过对式(3)的分析,可以归纳出逻辑曲线有两大显著特点:

(1) 当 $N \leq k$ 时,此为初始阶段的曲线, $dN/dt \approx bkN$,对此求积分,得到曲线初始阶段知识量的变化规律,近似为: $N = (k/a)e^{kbt}$,即指数增长模型。此阶段,两个增长模型并不矛盾,并且趋于一致。

(2) 知识累积趋于饱和状态。在 $t \in (0, \infty)$ 的区域内,始终有 $dN/dt > 0$,即知识量是单调递增函数。当 $t < \ln a/(kb)$ 时,知识增长速度加快,这是快速成长期;当 $t > \ln a/(kb)$ 时,知识增长速度减慢,这是成熟阶段;当 $t \rightarrow \infty$ 时,知识增长速度 dN/dt 趋于 0,知识量 N 趋于

常数 k ,即达到了饱和状态值。

逻辑增长模型描述了学科发展全过程的规律,但它预言了学科发展的极限状态,这表明用逻辑曲线描述知识总量增长情况具有局限性。

鉴于此,普赖斯文和纳里莫夫将两个模型进行综合,生成“普赖斯文—纳里莫夫循环曲线(简称 P. H 循环曲线)”增长模型。他们对科学发展的历史统计分析后认为:学科发展初期,知识累积量呈指数规律增长;进入相对成熟期,学科发展增长率变小,并逐渐趋于零。这预示着科学研究突变的来临。实现了科技飞跃后的学科又重复逻辑曲线增长规律,呈现螺旋式前进、上升趋势。

但是,P. H 循环曲线简单地认为,在整个时间长河,科学发展周期是不变的,这也违背了科学发展的常理。因此上述两种增长模型,不适合用来描述科学发展整个过程中的知识增长规律,只能准确地描述某些阶段的知识量增长规律。实际上,知识增长循环的周期在整个时间长河中逐渐递减。

2 知识老化与知识折旧率

随着科学技术的迅猛发展,新的知识不断产生,旧的知识不断被取代,知识的这种新陈代谢已成为现代科学技术发展的一个显著特点。这表明:在测度知识存量时,不能将历史上产生的所有知识均计算在内,而必须考虑知识的老化问题。因此,如何确定知识老化率是测度知识存量的一个关键问题。

2.1 知识的老化现象

知识老化是知识使用价值时效性的客观表现。正如美国情报学家戈斯内尔(C. P. Gosnell)所提出的:“随着时间的推移,一切知识或其相应的载体会逐渐失去原有的价值。”

2.1.1 知识老化的表现形式

知识老化主要有以下几种情形:①原有的知识后来被证明是不可靠的甚至是错误的;②由于新知识的产生,原有的知识不能再用于生产,被新知识取代;③对某一组织而言,其专有知识被其它组织所了解。

2.1.2 影响知识老化的因素

知识老化是一个复杂的动态过程,受多种因素的影响。

(1) 与知识的类别有关。知识老化随知识类别的不同而不同。一般来说,基础科学知识包括重要理论、实验方法等,比应用技术知识老化的速度慢,如物理学基本定律的老化速度远低于应用电子学的老化速度。以客观事实和实验数据为主的知识相对稳定,老化速度较慢,如地质学、植物学等。即使是应用技术,其老化速度亦因技术领域不同而不同,如信息技术的老化速度快于机械技术。

(2) 与知识所处的发展阶段有关。知识的发展一

般要经历诞生、发展、相对成熟等阶段。当知识处于诞生和发展的初期,知识尚不成熟,吸收新知识多,更新快,老化速度也较快;当知识进入相对成熟期后,知识的增长速度放慢,其老化速度相应变慢。因此知识所处的发展阶段不同,知识老化速度亦不同。

(3)与知识的资源禀赋有关。经济发达的国家或地区,资源禀赋优越,知识生产的投入多,知识更新快,知识老化速度快;经济落后的国家或地区,资源禀赋较差,知识生产的投入少,知识更新速度慢,知识老化速度也慢。因此一个国家或地区知识老化的速度与其经济发展水平和整体实力有关。

(4)与国家的知识发展战略有关。如果一个国家或地区意识到知识对经济发展的重要性,将发展科学技术、提高人才素质、促进技术创新放在战略高度来考虑,将有限的资源重点放在知识生产或知识引进上,则该国家或地区的知识水平可能出现跳跃性发展,知识老化的速度可能成倍提高^[3]。

由此可见,如何正确认识知识老化规律、合理确定知识老化率,是测度知识存量的关键问题。

2.2 知识老化率或折旧率的确定

知识老化率或折旧率实质上是指知识利用价值的耗蚀速度。知识存量测定的关键在于如何相对准确地测定知识折旧率。澳大利亚对折旧率取值为 10%,其它国家有的取值 15%,究竟取值多少较符合实际,取值的依据又是什么,这些问题有待人们作进一步的研究^[4]。

本文认为,知识老化必须具体问题具体分析,不能简单地一概而论。结合上面的分析,本文从以下几个方面进行分析:

2.2.1 经济系统中的知识老化与折旧问题

对进入经济系统的知识而言,其寿命取决于市场。当某一知识不能为知识持有者带来额外收益,失去其利用价值,被新的知识所替代时,该知识的经济寿命就结束了。

对经济系统的知识折旧率,可以采用以下两种方法测算:

(1)法定年限法。有些知识,知识产权法律规定了保护期或知识转让合同规定了使用年限。对这类知识往往以法定年限为其经济寿命,如发明专利为 20 年。若某知识的法定年限为 n 年,则其折旧率 δ 为:

$$\delta = \frac{1}{n} \tag{4}$$

若知识更新的时间短于法定时间,则按此方法计算的折旧率偏低。

(2)统计法。对没有规定法定年限的知识,可以通过统计方法得出这类知识更新的周期,由此确定该知识的折旧率。

2.2.2 科学文献知识老化与折旧问题

(1)科技文献老化的负指数模型。英国著名文献

计量学家布鲁克斯提出,科技文献被引用的数量服从负指数规律(见图 2)。此后,奥利费等人证实了这一规律的存在。该模型为:

$$C(t) = Ke^{-\alpha t} \tag{5}$$

其中, t 为被引用文献的出版年龄; $C(t)$ 为 t 年的引文数量; e 为自然对数的底数; K 为随学科而异的常数; α 为文献的老化率。

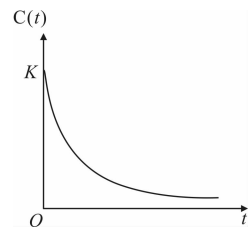


图 2 负指数模型函数曲线

(2)伯顿-开普勒公式的发明者美国图书馆员伯顿和物理学家开普勒,选择物理、化学、机械工程等 9 种学科的有关文献进行了引文统计分析,他们按引文年龄由小到大的顺序,以 10 年为单位统计了这 9 个学科的引文积累百分比,并根据经验数据得出了下列方程式:

$$y = 1 - \left(\frac{a}{e^t} + \frac{b}{e^{2t}} \right), \text{ 且 } a + b = 1 \tag{6}$$

其中, y 为经过时间 t 后该学科尚在利用的文献的相对数量; t 为被引文献出版年龄(以 10 年为单位); a 、 b 为因学科而异的常数。

其函数曲线如图 3 所示。

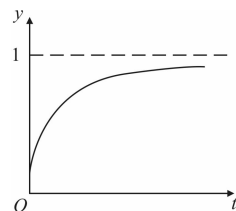


图 3 伯顿-开普勒函数曲线

莫德列夫研究伯顿-开普勒方程后,发现其拟合效果不理想,遂提出了如下改进公式:

$$y = 1 - \left(\frac{a}{e^{t-0.1}} + \frac{b}{e^{2t-0.2}} \right) \tag{7}$$

其中, $a = 3.4596 - 4.1447y_i$; $b = 1 - a$; y_i 为实测 10 年累积引文相对比率。

所谓实测 10 年累积引文相对比率,是指选择该学科领域有代表性的重要杂志,统计这些杂志一年内所发表的论文所有引文的数量——被引文量,然后分别统计出引用不同年份的论文数量,计算出该年份相对于被引文量的比率、累积最新 10 年的比率所得到的值。

由式(7)可求出文献老化率:

$$\delta = \frac{a}{e^{t-0.1}} + \frac{b}{e^{2t-0.2}} \tag{8}$$

3 知识存量测度

3.1 利用科技文献测度知识存量的可行性

由以上分析可知,知识量和科技文献的增长紧密相关,具有很大的一致性。科技文献是人类知识存储和表现的主要形式。特别是文献期刊,及时反映学科的发展动向,是确定创新成果是否具有优先权的最好依据。所以,科技文献在很大程度上反映了知识存量的基本状况,是衡量知识存量的重要尺度。科技文献数据时间序列长、易于统计,以其作为知识存量的测度指标简便、可行。

3.2 基于科技文献的知识存量测度

3.2.1 科技文献时效性

科技文献所记载的知识随着时间的推移,价值越来越小,这就是科技文献的老化现象。在以其衡量知识存量时,须加以剔除。关于科技文献的老化问题,本文第 2 部分已经作了专门分析。

3.2.2 科技文献影响指数

从单篇文章看,论文质量差异是难以估量的,但从总体上看,还是可以在一定程度上进行衡量的。

对于科技文献的质量和影响,这里以引文分析为基础,采用影响指数展开分析。用下式表示:

$$\xi_i = \frac{n_i}{N_i} \quad (9)$$

式中, ξ_i 为某学科文献的影响指数; n_i 为在某年度内某学科论文的被引用数量; N_i 为在某年度内所有学科论文被引用总量。

3.2.3 基于科技文献的知识存量测度公式

有 m 个学科,考虑科技文献的时效性和影响指数,总的知识存量为:

$$S_t = \sum_{i=1}^m S_{t,i} = \sum_{i=1}^m \{(1 - \delta_i) S_{t-1,i} + \xi_{t,i} \cdot R_{t,i}\} \quad (10)$$

式中 S_t 为 t 时刻知识总的存量; $S_{t,i}$ 为 t 时刻的知识存量(学科 i); δ_i 为知识折旧率(学科 i); $S_{t-1,i}$ 为 $t-1$ 时刻的知识存量(学科 i); $\xi_{t,i}$ 为 t 时刻的文献影响指数(学科 i); $R_{t,i}$ 为 t 时刻新增的知识存量(学科 i); i 为学科数, $i = 1, 2, \dots, m$ 。

从式(10)的表面看,计算量随着时间的推移,越来越大,但是,原有的文献资料经过较长一段时间以后,其值几乎接近于零。因此,按一定精度要求,只需取若干年的数值,这样就可大大减少计算量^[5]。

参考文献:

- [1] 苏屹,李柏洲,喻登科. 区域创新系统知识存量的测度与公平性研究[J]. 中国软科学, 2012(5): 157-174.
- [2] 吴新文,揭建群,赵军. 基于信息功能和商品功能的知识存量测度研究[J]. 科技进步与对策, 2010(24): 128-130.
- [3] 杨鹏,许晓雯,蔡虹,等. 我国区域 R&D 知识存量与 GDP 的实证检验[J]. 科学学与科学技术管理, 2005, 26(12): 23-26.
- [4] 李顺才,邹珊刚,苏子仪,等. 一种基于永续盘存的知识存量测度改进模型[J]. 科学学与科学技术管理, 2003, 24(9): 13-15.
- [5] 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.

(责任编辑:陈晓峰)

Knowledge Measurement Based on Knowledge Growth Pattern

Wu Xinwen¹, Xiong Yonghao¹, Zhao Sa², Liang Jingyan¹, Ma Jing³, Huang Xiaoling³

(1. Hubei Institute of S&T Information, Wuhan 430071, China;

2. China University of Geosciences, Wuhan 430074, China;

3. Hubei University of Economics, Wuhan 430205, China)

Abstract: This paper analyzes two classic models; the pattern for knowledge exponential growth and logic growth and probes into the phenomenon of knowledge depreciation and the influencing factors. Therefore, the paper defines the calculation method of knowledge depreciation rate and makes research on the knowledge measurement pattern based on scientific articles.

Key Words: Knowledge Measurement; Exponential Growth; Logic Growth; Depreciation Rate of Knowledge; Depreciation Factors