

时间因素视角下科研人员评价的 Pt 指数研究

韩毅 夏慧

摘要 P 指数是在克服 h 指数缺乏区分度、仅关注 h 核内元素和操作繁琐等缺陷的基础上发展起来的新型科技评价指标,但不考虑时间因素的 P 指数,难以克服相同引文量不同引文时间窗文献和老化速度较慢文献的评价误差。本研究在分析时间因素的稀释效应后,构建了评价相同引文量不同引文时间窗文献的 P_{t_1} 指数,在分析时间因素的凝聚效应后构建了评价老化速度较慢文献的 P_{t_2} 指数。选择图书情报与文献学领域 h 指数排名前 100 位的学者为样本,以 CSSCI 收录的文献为对象,采集样本对象 1998—2013 年的发文量和引文量数据,计算样本对象的 h 指数、P 指数、 P_{t_1} 指数和 P_{t_2} 指数。结果表明,两个 Pt 指数是高效的综合质量因素与数量因素的评价指标;相较于 h 指数,两个 Pt 指数均具有较高的区分度和灵敏性;相较于 P 指数, P_{t_1} 指数具有识别发文时长较短的潜在高价值文献的能力, P_{t_2} 指数具有识别经典高价值文献的能力。表 9。参考文献 25。

关键词 科研人员评价 时间稀释效应 时间凝聚效应 Pt 指数

分类号 G250.2

Pt Indices: Evaluating Researchers Under the Context of Time

HAN Yi & XIA Hui

ABSTRACT

This paper, taking account of the time factor, intends to reconstruct P-index, a new scientific and technological evaluation index which could overcome such problems of h-index as lack of discrimination, mere focus on core elements and complex operations yet still encountering difficulty, due to its neglect of the time effect, in reducing evaluation errors under the context of evaluating literature with the same citation amounts but different citation time window or slow-aging literature. Time, a historical selective factor, is taken as a regulating factor to improve P-index owing to its certain selectively filtering function upon the researchers' scientific products and the corresponding general classification of the quality levels of researchers' scientific products. For the researcher's scientific products with different publication age, the publication age at the evaluation point is the dilution factor to its quality factor, namely citation number, which is referred to as dilution effects of the time factor. Consequently, P_{t_1} index, with the publication age as a regulating factor, is constructed to evaluate the scientific performance of the researchers and to eliminate the cumulative effect of the time factor. For the researcher's products collection, the citation distribution characteristics of different documents reflect their contribution differences to researcher's prestige, and the time for the long-time cited literatures is an agglomerative factor to the citation numbers, which is called agglomerative effect of time factor. P_{t_2} index is thus constructed, with the half life of one

通信作者:韩毅,Email:Hanyi72@swu.edu.cn,ORCID:0000-0001-7021-3229 (Correspondence should be addressed to HAN Yi,Email:Hanyi72@swu.edu.cn,ORCID:0000-0001-7021-3229)

discipline as the boundary line to weigh the quality factor of cited literature out of the half life, to evaluate the researcher's performance and highlight the contribution rate of high-quality literature.

In this research, the top 100 scholars of Library and Information Science in China are sorted by h-index as a sample and the data about citing and cited literature are retrieved in the database of Chinese Social Sciences Citation Index (CSCSI) from 1998 to 2013. The half-life of about the sample discipline is calculated, based on the Burton-Kebler equation, as about 4 years, and the parameter b as about 0.29 based on the exponential model of citation distribution. Citation rate (C/N), weighted Pt_1 citation amount, weighted Pt_2 citation amount, weighted Pt_1 citation rate ($CW1/N$), weighted Pt_2 citation rate ($CW2/N$), h index, P index, Pt_1 index and Pt_2 index are all calculated. Also calculated are some correlation coefficients such as between papers number, citations number, citation rate and other indices, weighted Pt_1 citations amount, weighted Pt_1 citation rate, Pt_1 index and other indices, weighted Pt_2 citations amount, weighted Pt_2 citation rate, Pt_1 index and other indices.

It is clarified by the research results that compared with h-index, the two Pt indices, with higher discrimination and sensitivity, are effective and comprehensive evaluation index of quality factor and quantity factor; compared with P -index, Pt_1 -index has the potentiality to identify would-be high-value literature with short publication age, and Pt_2 index has the ability to identify classic high-value literature. There is still much to improve in this paper such as the necessity of the further test of the conclusion out of the one-sidedness of the concerned sample discipline and the possibly increasing difficulty of calculation as a result of the modified formula which is more complex than the original one. 9 tabs. 25 refs.

KEY WORDS

Researchers evaluation. Dilute effect of time. Agglomerative effect of time. Pt indices.

0 问题提出

科技评价目前主要有两种评价方法:以质性评估为核心的同行评议和以量化评估为手段的文献计量法^[1]。同行评议可以充分发挥同行专家的经验 and 智慧,适用于需要对内容进行评价或某些因素难以客观量化情况下的成果评价,但评价结果具有一定的主观性,有时甚至带有较大的偏见^[2]。文献计量方法是利用科技文献发表时和发表后所产生的效应,通过设计评价指标对科技成果进行定量分析,但这种评价更多是关注科技成果的外在特征^[3]。

科研人员评价作为科技评价的重要组成部分,以文献计量方法进行评价是其中的重要方式。早期评价主要是单一指标,如发文量、引用量等,如何设计更高层面的综合性指标成为计

量学研究的一大难题,Hirsch 提出的 h 指数为该难题的解决提供了一种思路^[4]。但 h 指数缺乏区分度,操作繁琐,迫使人们探索新的综合性指数, P 指数的提出在一定程度上弥补了 h 指数缺乏区分度的缺陷^[5]。但 P 指数并没有考虑与评价对象有关的时间因素,如一篇发表 10 年被引用 100 次的文章和发表两年被引用 100 次的文章,其影响力应该是不同的,再如两篇总被引量同为 100 次的文章,其中一篇被引量逐年递减,而另一篇被引量逐年递增,其影响力也应该是不同的。因此,构建一个更高层面的,能综合质量、数量,体现引文时间分布特性的综合性指标,无论在理论上还是实践上都具有重要意义。

1 相关研究工作

关于综合性评价指标的研究最早可追溯到

Garfield 基于科学引文索引数据库创立的影响因子指标,它综合考察期刊引文量与发文量的关系来计算篇均引文量,这种线性关系的表述虽然简单,但对于以指数变化为特征的引文分布的描述并不恰当;1978年,Lindsey 通过综合分析发现,可用修正质量比(Corrected Quality Ratio, CQ)指标代替平均引用率作为期刊质量的综合评价指标, $CQ = (C/N) (CN)^{1/2} = (C^3/N)^{1/2}$ (C 指引文量, N 指发文量),能在一定程度上达到 C、N、和 C/N 之间的均衡^[6]。

2005年,Hirsch 提出的以科研人员绩效为评价对象的 h 指数,也是综合了发文量与引文量的综合性指标,但其操作过程需要对评价对象进行排序处理,不如 CQ 计算那么方便。既然 h 指数与发文量和引文量都有关系,从数学上一定能找出其经验公式。实践证明 h 指数与 $x^a/(a+1)N^{1/(a+1)}$ 之间存在强相关($x=C/N$,其中 C 为引文量, N 为发文量)^[7]。在不同的引文时间窗下, a 值在 1.5—2.0 之间变化;当 a=2 时, $h = cN^{1/3}x^{2/3}$ (c 为常量, N 为发文量, x 为引用率),将 h 指数与 $N^{1/3}x^{2/3}$ 进行数据拟合,拟合度 $R^2 = 0.95$ ^[8]。Csajbok 等通过实证数据得出 $h=0.932(C^2/N)^{1/3}$,其中 h 指数与 $(C^2/N)^{1/3}$ 两者的数据拟合度 $R^2 = 0.988$ ^[9]。

2010年,Prathap 认为 $(C^2/N)^{1/3}$ 内涵丰富,可看成一个新的综合指标,具有模拟 h 指数的功能,他用 hm 来代表 $(C^2/N)^{1/3}$,并利用该指标分析全球主要国家的科学产出贡献,评价三个学术机构的科学影响力^[10]。随后他又用 P (Prestige) 指数来代替 hm,即 $P = hm = (C(C/N))^{1/3}$,并用 P 指数对 100 位经济学家进行排序,实践证明 P 指数与 h 指数存在极大相关性,且比 h 指数更通用^[11]。P 指数作为科研评价的综合性指标,反映评估对象在特定系统中的位置属性,表征评估对象在特定方面的潜在做功能力,其计算公式与电力学中的电能和动力学中的动能公式形式上完全相同,因而其表征的物理学意义也应该相似,都反映了特定对象的潜在运动能力^[12]。

自 P 指数提出后,学者从多个方面展开研究,如开发测度国家科研能力的 iCE 二维地图^[13-14],探讨合著论文科学家 P 指数的计算方法^[15],类比于热力学深入分析 P 指数的功能^[16],并从整体视角探讨 P 指数在绩效评估中的作用机理^[17-18],但更多的研究是应用 P 指数进行测度评价,如对期刊的评价^[13,19-20],对图书馆馆藏的评价^[5],对科研人员的评价^[6,11]。研究表明,P 指数简洁、易计算、区分度好,支持动态变化排名,与多个关键评价指标相关性好,且在优秀学术期刊识别方面具有较好的可靠性。

h 指数作为一个科研人员评价指标,综合考察了评价对象的质量因素与数量因素,但在评价过程中科研人员有 75% 的引文都被排除在外^[21]。事实上,科研人员学术生涯中的一举一动(每发表一篇文章,每一次被引用)都会对其学术地位产生影响,P 指数充分考虑了科研人员每一篇文章、每一次被引用,并以几何平均数来平衡引文量与发文量的关系,应该是更优的评价指标。然而在实际评价过程中,从评价时间点回溯,作者在不同时间窗发表的文献,文献在发表后的引用分布差异,将深刻地影响科研人员的学术地位,因而需要考虑时间因素对 P 指数评价功能的影响,本文正是基于这样的视角来探讨 Pt 指数的构建问题。

2 方法与数据

2.1 时间因素视角下的 Pt 指数构建

时间对科学评价的影响是多方面的。首先,随着时间演进,某个时点发表文献的累计引文量有增无减,对于有相同引文数量,但在评估时点其发表年龄不同的文献来说,显然发表年龄越短的文献,其学术影响力更大,但 P 指数无法反映出该情形下时间因素的影响作用;如果不考虑时间因素,即使评价对象不再发表新的文献,其 P 指数也恒定不变,这样的评价对于科研人员没有激励作用。其次,随着时间推移,某个时点发表的文献逐渐老化,引用峰值年后,文

献引文量随着时间递减,影响力小、价值低的文献其引文量很快就衰减到 0,而影响力大、价值高的文献老化速度慢,引用峰值年后,其引文数量衰减缓慢,被引用时间长,因此如果不考虑时间因素,简单的 P 指数也难以突显这部分高价值文献的影响力。对于第一种情况,时间因素是一种稀释效应,相同引文数量的文献发表时间越长,其稀释效应越明显,因而在评价上需要考虑时间因素稀释效应的影响;对于第二种情况,仿佛时间被收缩一样,时间因素是一种凝聚效应,文献老化速度越慢,文献学术影响力的凝聚效应越明显,因而在评价上需要突显出时间凝聚效应的影响。

欲以指数 $P = (C^2/N)^{1/3}$ 来评价科研人员的学术影响力,需要统计科研人员的发文量与引文量两个参数。为了数据计算方便,发文量 N 是以独著和排名第一的发文数简单加和,没有考虑其他发文情况;由于时间因素的影响,引文

量 C 不能是简单的算术加和,而应是考虑时间因素的加权值,即根据时间因素加权计算每篇文献的引文量,再将所有加权引文量求和得到总引文量。为了区别于不考虑时间因素的 P 指数,本文把考虑时间效应的指数称之为 P_t 指数。

2.1.1 聚焦稀释效应的 P_{t_1} 指数

对于有一定利用价值的文献来说,发表时间越长的文献获得的累计引文越多。但对于相同引文数的评价对象来说,文献的总引文量可根据发文时间进行简单平均作为考虑时间因素稀释效应后的加权引文量,即:

$$CW_i = \frac{C_i}{T_i}$$

CW (Citation-Weighted) 表示评价时点或统计时点的加权引文量, CW_i 表示第 i 篇文献的加权引文量, C_i 表示第 i 篇文献的统计时点引文量, T_i 表示统计时点第 i 篇文献的发文时长或文献年龄。如此可得到考虑时间稀释效应的 P_{t_1} 指数:

$$P_{t_1} = \left(CW * \frac{CW}{N} \right)^{\frac{1}{3}} = \left(\left(\frac{\sum_{i=1}^{i=N} C_i}{\sum_{i=1}^{i=N} T_i} \right) * \left(\frac{\left(\frac{\sum_{i=1}^{i=N} C_i}{\sum_{i=1}^{i=N} T_i} \right)}{N} \right) \right)^{\frac{1}{3}}$$

之所以选择统计时点的发文时长作为调节因子计算平均加权引文数,一方面可识别出发表时长相同文献中的高被引文献,另一方面也可以体现出引文量相同文献的时间稀释效应。

2.1.2 聚焦凝聚效应的 P_{t_2} 指数

在学科发展过程中,影响力越大价值越高的文献老化速度越慢,某个科研人员的这种高价值文献越多,其在学科发展过程中的地位越重要,因而在评价过程需要突出这部分高价值文献的作用。这样就引发两个问题:一是如何识别出这些高价值文献,二是如何为这些高价值文献进行引文量加权。

对一个学科或某一文献集合来说,在统计时点上其引用文献的时间分布十分广泛,半衰期较好地刻画了这种引文时间分布特征。据此,本文以学科半衰期为阈值来识别高价值文献,即某一文献在学科半衰期外还有引用发生,

则认为该文献价值较高,应对其半衰期外产生的引文赋予更高权重;而对半衰期内发生的引用按简单加和处理。

如何确定半衰期外的文献引文量调节因子呢? 本文采用引文量的指数分布函数来进行处理。一篇发表 a 年的文献,其引文量分布可用下面的指数函数来表达:

$$C(a) = b_0 * e^{-b_1 a}$$

公式中 b_0, b_1 为函数参数,不同学科的参数不同, $b_0 > 0, b_1 < 0, a$ 为文献年龄。

$C(a)$ 作为文献年龄为 a 的引文量,如果文献的价值低,则 $C(a-1)$ 的值会迅速减小,因而可使用 $\frac{C(a-1)}{C(a)} = e^{-b_1}$ 作为高价值文献半衰期外引文量的调节因子。某一文献在半衰期后的引文加权值为:

$$CW(a) = C * e^{-(a-t_h)b_1} \quad (a > t_h)$$

公式中, a 为某篇文献的文献年龄, C 为文献年龄为 a 时的引文量, t_h 为学科半衰期, CW

(a) 为文献年龄为 a 时的加权引文量。

如此可得到考虑时间凝聚效应的 Pt_2 指数:

$$Pt_2 = \left(CW * \left(\frac{CW}{N} \right) \right)^{\frac{1}{3}}$$

$$CW = \begin{cases} \sum_{k=1}^{k=N} \left(\sum_{a=t_h}^{k_a} C_{k_a} + \sum_{a=0}^{k_a-t_h} (C_{k_a} * e^{-(k_a-t_h)b_1}) \right) & (k_a - t_h > 0) \\ \sum_{k=1}^{k=N} \left(\sum_{a=0}^{k_a} C_{k_a} \right) & (k_a - t_h \leq 0) \end{cases}$$

公式中, CW 为评价对象全部发文的加权引文量总和, k 表示评价对象的第 k 篇发文, k_a 为第 k 篇发文的发文时长, C_{k_a} 为第 k 篇发文在发表后第 a 年的引文量, t_h 为学科半衰期, b_1 为基于引文量指数模型估计出的学科参数。也就是说, 对于一篇文献来说, 如果其发文时长在半衰期范围以内, 则其引文量直接相加即可; 如果其发文时长大于半衰期且半衰期外还有引文的话, 则需突出半衰期外的部分而采用指数加权引文量, 半衰期内的仍采用简单求和。

选择其中 h 指数大于 5 的科研人员 85 位, h 指数等于 5 中发文量靠前的 15 位科研人员, 总计 100 人作为研究对象。

学科整体被引数据收集以 CSSCI 数据库收录的该领域期刊为基础进行, 分别检索收录期刊被引数据, 然后基于被引年代统计该期刊某一年被引文献的引文分布。

以评价对象姓名检索该作者在 1998—2013 年 CSSCI 收录的所有文献, 记录各作者所发表文献的篇名及发表时间, 整理成 txt 文档; 然后将 txt 文档导入自编程序, 便可得到该作者每篇发文的被引数据、所有期刊的被引数据, 以及所有科研人员的发文与被引数据。

2.2 数据来源

目前, 我国可用于引文分析的引文数据库种类较多。其中, 中国科学引文数据库 (CSCD) 和中国科技论文与引文数据库 (CSTPCD) 收录自然科学文献, 中文社会科学引文索引 (CSSCI) 收录社会科学文献, 中国引文数据库 (CCD) 和中文科技期刊引文数据库 (VIP) 则同时包含自然科学和社会科学的文献。为了验证前面构建 Pt 指数的有效性, 本文以熟悉的图书情报与文献学领域为样本, 选择 CSSCI 为目标数据来源库。

在数据收集过程中需要说明以下几点: ①数据引文时间窗为 1998—2013 年, 因而本文仅对 1998—2013 年间科研人员的学术地位进行评价, 并不完全代表科研人员在整个历史长河中的学术地位; ②数据仅限于科研人员 1998—2013 年在 CSSCI 来源期刊中的发文, 不包括非 CSSCI 来源期刊上发表文献的被引量; ③CSSCI 数据库的来源文献不包含图书, 为了使科研人员学术影响力评价更符合实际, 补充检索样本对象图书的被引情况; ④由于 CSSCI 对检索式的长度有所限制, 为了消除自编程序因长度原因产生的相同引文数据, 对这部分文献进行重新检索校正; ⑤对同名作者的数据以作者机构或作者地区加以区分, 对于同一作者先后在不同机构工作的情况, 分别检索其不同工作机构的文献再汇总。

本文的数据采集时间为 2015 年 3 月 10 日, 由于数据入库延迟的原因, 主要采集 CSSCI 数据库中 1998—2013 年的引文数据, 故引文数据时间窗为 1998—2013 年。在确定图书情报与文献学领域评价对象时, 主要参考王志军^[6]、朱惠^[22]、邱均平^[23-24]、丁楠^[25] 等学者的研究成果, 整理出其中提及的图书情报与文献学领域 256 位科研人员, 分别检索其发文和引文数据,

3 研究结果

整理数据得到图书情报与文献学领域 1998—2013 年的引文数据如表 1 所示。表格横

行为被引文献的发表时间,纵栏为被引文献的被引时间,单元格数据是指引用年文献对发文年文献的引用情况,如表格中发文时间为 1998 与被引时间为 2013 相交的单元格中数字 53,表示的意义是 2013 年引用 1998 年发表的文献共 53 次。

表 1 图书情报与文献学领域 1998—2013 年的引文数据分布

被引时间 \ 发文时间	1998 前	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
1998	3 751	180	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1999	3 028	1 381	247	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2000	1 256	1 387	2 081	244	0	0	0	0	0	0	0	0	0	0	0	0	0
2001	1 157	780	1 557	1 933	300	1	0	0	0	0	0	0	0	0	0	0	0
2002	803	566	1 029	2 147	2 361	277	0	0	0	0	0	0	0	0	0	0	0
2003	856	469	721	1 603	2 383	2 354	322	0	0	0	0	0	0	0	0	0	0
2004	866	398	598	1 155	1 864	2 965	2 576	267	0	0	0	0	0	0	0	0	0
2005	621	276	378	820	1 170	1 928	2 644	2 386	360	0	0	0	0	0	0	0	0
2006	566	213	268	575	779	1 270	1 768	2 376	2 393	371	0	0	0	0	0	0	0
2007	483	140	172	413	507	861	1 151	1 651	2 309	1 969	270	0	0	0	0	0	0
2008	403	132	177	366	505	677	1 030	1 462	2 122	2 859	2 338	322	1	0	0	0	0
2009	481	138	150	376	406	616	831	1 261	1 746	2 330	2 829	2 418	319	5	0	0	0
2010	442	121	138	286	340	470	621	867	1 296	1 587	1 982	2 516	2 420	310	1	0	0
2011	315	80	78	154	179	228	330	459	685	962	1 046	1 410	1 821	1 267	60	0	0
2012	439	97	103	212	240	325	440	598	777	1 061	1 254	1 565	2 293	2 736	2 232	369	0
2013	311	53	49	115	135	191	234	334	446	579	750	840	1 280	1 590	1 890	1 407	170

3.1 图书情报与文献学领域的引用半衰期

基于上表的引用数据计算出 2004—2013 年各年最新的 10%、30%、50%、70%、90% 引用文献的最大文献年龄(即在多长时间内发表),结果如表 2 所示。表格中横行为引用年,纵栏为最新文献的利用率,单元格数据为引用文献对应最新引用百分比的最大引用年龄,如表格中引用年为 2013 最新文献利用率为 30% 和 50% 的单元格中数字 2.81 和 4.10,其涵义为:2013 年

引用的文献中,最新的 30% 是在 2.81 年内发表的,2013 年引用的文献中最新的 50% 是在 4.10 年内发表的。之所以只选择 2004—2013 年的被引数据进行分析,是因为 2003 年之前各年最新的 70%、90% 被引文献大都为 1998 年之前某一年至引用年发表,由于数据库原因,无法统计出 1998 年前的引文年代分布,因而本文对引用 1998 年之前发表文献的次数合计为“1998 年之前”,不能准确获得其最大文献年龄。

表 2 图书情报与文献学领域 2004—2013 年最新文献引用率的最大文献年龄

引用年 t \ 最新文献 利用率 u (%)	2013	2012	2011	2010	2009	2008	2007	2006	2005	2004
10	1.62	1.50	1.67	1.43	1.44	1.39	1.37	1.44	1.29	1.31
30	2.81	2.67	2.77	2.51	2.51	2.37	2.32	2.17	2.16	2.12
50	4.10	3.89	3.99	3.73	3.59	3.32	3.25	3.09	3.96	3.84
70	6.10	5.90	5.78	5.43	5.07	4.71	4.65	4.39	4.08	4.90
90	10.22	9.87	9.56	8.97	8.40	7.68	7.52	6.98	6.57	6.49

利用 SPSS 18.0 以巴尔顿—凯普勒 (Burton-Kebler) 老化方程为基础进行非线性回归分析, 得到各年回归结果如表 3 所示。根据回归结果可知, 各年回归模型参数估计值标准误差均较小, 说明各年参数值置信度较高; 各年模

型拟合度均在 0.85 及以上, 说明巴尔顿—凯普勒老化模型对数据的可解释度较高。对 2004—2013 年半衰期求均值可得到图书情报文献学领域半衰期为 3.7 年, 为方便计算取整为 4 年。

表 3 图书情报与文献学领域 2004—2013 年巴—凯老化方程回归结果

统计年	Parameter 参数值		Parameter Std. error 参数标准误差		Regression 回归平方和	Residual 残差平方和	模型拟合度 R ²	Half-time 半衰期
	a	b	a	b				
2013	0.252	0.748	0.413	0.567	1.611	0.039	0.902 5	0.406
2012	0.179	0.821	0.389	0.527	1.615	0.035	0.912 5	0.387
2011	0.199	0.801	0.463	0.636	1.604	0.046	0.885	0.392
2010	0.067	0.933	0.407	0.546	1.613	0.037	0.907 5	0.361
2009	0.001	0.999	0.458	0.614	1.607	0.043	0.892 5	0.346
2008	0.036	0.964	0.487	0.645	1.605	0.045	0.887 5	0.354
2007	0.006	0.994	0.486	0.642	1.606	0.044	0.89	0.347
2006	-0.066	1.066	0.551	0.724	1.599	0.051	0.872 5	0.333
2005	0.04	0.996	0.576	0.759	1.589	0.061	0.847 5	0.373
2004	0.08	0.92	0.505	0.667	1.599	0.051	0.872 5	0.364

3.2 图书情报与文献学领域的引文权重

整理图书情报与文献学领域 2008—2013 年各年被引文献的年龄分布如表 4 所示。表中横行为文献引用时间, 纵栏为被引文献的文献年龄, 单元格数据是引用年引用的不同文献年龄

的文献量。如, 引用年 2013 与文献年龄 5 交叉形成的单元格中数值 1 280 表示的含义是, 2013 年文献年龄为 5 年的文献被引 1 280 次, 或者说 1 280 次是引用发表 5 年的文献。

表4 图书情报与文献学领域 2008—2013 年各年被引文献的年龄分布

引用年 文献年龄 a	2013	2012	2011	2010	2009	2008
1	170	369	60	310	319	322
2	1 407	2 232	1 267	2 420	2 418	2 338
3	1 890	2 736	1 821	2 516	2 829	2 859
4	1 590	2 293	1 410	1 982	2 330	2 122
5	1 280	1 565	1 046	1 587	1 746	1 462
6	840	1 254	962	1 296	1 261	1 030
7	750	1 061	685	867	831	677
8	579	777	459	621	616	505
9	446	598	330	470	406	366
10	334	440	228	340	376	177
11	234	325	179	286	150	132
12	191	240	154	138	138	
13	135	212	78	121		
14	115	103	80			
15	49	97				
16	53					

为了计算凝聚效应的 Pt_2 指数半衰期外引文的权重,将上述图书情报与文献学各年被引文献的年龄分布进行指数分布模型的曲线估计,结果如表 5 所示。之所以不对 2008 年之前各年数据进行曲线估计,是因为 2008 年之前各

年数据量太少,可能会影响模型拟合准确度。此外,各年文献年龄为 1 的被引量未参与模型估计,因为文献发表第一年一般不会大量引用,若加入文献年龄为 1 的文献被引量进行估计,可能会影响曲线估计效果及模型拟合度。

表5 图书情报与文献学 2008—2013 各年文献年龄与被引量的曲线估计结果

估计年	模型概况						系数		系数显著性		
	拟合度 R^2	回归 平方和	残差 平方和	自由度		F 值	模型显著 性 Sig.	b_0	b_1	b_0	b_1
				df1	df2						
2013	0.967	19.727	0.678	1	13	378.05	0.000	3190.192	-0.265	0.000	0.000
2012	0.976	16.102	0.394	1	12	490.837	0.000	4485.396	-0.266	0.000	0.000
2011	0.960	13.786	0.568	1	11	267.03	0.000	2894.966	-0.275	0.000	0.000
2010	0.975	12.251	0.308	1	10	397.547	0.000	4574.198	-0.293	0.000	0.000
2009	0.957	11.246	0.507	1	9	199.696	0.000	5246.874	-0.32	0.000	0.000
2008	0.963	10.102	0.392	1	8	205.911	0.000	5192.841	-0.35	0.000	0.000

根据上述曲线估计结果可知,图书情报与文献学领域 2008—2013 年被引文献与文献年龄指数回归系数分别为 -0.35、-0.32、-0.293、-0.275、-0.266、-0.265,各年指数分布模型拟合度均在 0.95 以上,拟合度优良,模型及模型系数估计均有较好显著性。对各年回归系数 b_1 求均值得到 -0.29,则时间凝聚效应的 Pt_2 指数半衰期外引文权重调节因子 e^{-b_1} 为 $e^{0.29}$ 。

3.3 图书情报与文献学领域的 Pt 指数

统计各科研人员的发文量(N)、引文量(C)

和引文率(C/N),分别计算 Pt_1 加权引文量(CW_1 ,基于公式 Pt_1 对各篇文章引文加权后求和)、 Pt_2 加权引文量(CW_2 ,基于公式 Pt_2 对各篇文章引文加权后求和)、 Pt_1 加权引文率(CW_1/N)、 Pt_2 加权引文率(CW_2/N)、h 指数、P 指数、 Pt_1 指数、 Pt_2 指数,以 h 指数为序得到如表 6 所示结果,此处仅列出 100 位科研人员中的前 20 位。

根据表 6 的计算结果可以看出,P、 Pt_1 、 Pt_2 指数比 h 指数具有更高的区分度, Pt_1 指数的变化范围较小,而 Pt_2 指数的波动幅度较大。由于 Pt_1 指数聚焦文献影响力的稀释效应,消除发文

表 6 图书情报与文献学领域 20 位科研人员的各参数计算值

序号	作者	N	C	C/N	CW_1	CW_2	CW_1/N	CW_2/N	h	P	Pt_1	Pt_2
1	张晓林	76	1211	15.93	142.38	2246.19	1.87	29.56	20	26.82	6.44	40.49
2	邱均平	219	1819	8.31	185.72	2813.66	0.85	12.85	20	24.72	5.40	33.06
3	蒋永福	84	958	11.40	103.64	1378.78	1.23	16.41	18	22.19	5.04	28.29
4	马费成	72	1126	15.64	109.8	1680.63	1.53	23.34	17	26.02	5.51	33.98
5	吴慰慈	51	821	16.10	75.48	1393.79	1.48	27.33	15	23.64	4.82	33.65
6	胡昌平	80	841	10.51	89.98	1264.41	1.12	15.81	15	20.68	4.66	27.14
7	王子舟	54	651	12.06	72	1107.87	1.33	20.52	15	19.87	4.58	28.33
8	黄宗忠	41	716	17.46	62.54	1093.97	1.53	26.68	14	23.21	4.57	30.79
9	肖希明	69	544	7.88	72.82	702.51	1.06	10.18	14	16.25	4.25	19.27
10	王知津	208	872	4.19	87.23	1338.24	0.42	6.43	14	15.40	3.32	20.50
11	范并思	54	839	15.54	97.49	1092.62	1.81	20.23	13	23.53	5.60	28.07
12	于良芝	35	575	16.43	68.94	771.67	1.97	22.05	13	21.14	5.14	25.72
13	程焕文	57	636	11.16	72.08	859.41	1.26	15.08	13	19.22	4.50	23.49
14	苏新宁	47	481	10.23	65.36	593.55	1.39	12.63	13	17.01	4.50	19.57
15	包昌火	30	649	21.63	71.04	1008.15	2.37	33.61	12	24.12	5.52	32.36
16	程亚男	38	364	9.58	31.27	561.33	0.82	14.77	12	15.16	2.95	20.24
17	柯平	94	465	4.95	58.53	711.39	0.62	7.57	12	13.20	3.32	17.53
18	陈传夫	62	432	6.97	55.97	560.69	0.90	9.04	11	14.44	3.70	17.18
19	黄俊贵	32	293	9.16	24.8	384.36	0.78	12.01	11	13.90	2.68	16.65
20	盛小平	56	346	6.18	35.61	430.33	0.64	7.68	11	12.88	2.83	14.90

时长的影响,对其引文加权处理时采用以发文时长为调节因子的算术平均方法,这就会使得加权后的引文总量小于原来的引文总量;而 Pt_2 指数聚焦文献影响力的凝聚效应,为了突出老化速度较慢的文献,对其引文加权处理时采用突显半衰期外的引文的方法,这就会使其引文量变大,从而使得两个 Pt 指数的变化区间具有明显差别。

表7为发文量、引文量、引文率与各指标之间的相关分析结果。从结果数据可以看出:科

研人员的发文量与其他计量指标并无强相关关系,而引文量却与 h 、 P 、 Pt_1 、 Pt_2 等评价指标具有较高相关性,说明决定科研人员学术影响力大小的是其发表文献有多大的利用价值,而不是发表文献的数量多少。引文率则与 P 、 Pt_1 、 Pt_2 指数具有较好相关性,但与 h 指数相关性不是太高,主要原因是 h 指数对各科研人员的区分度不高,很多引文率不同的科研人员具有相同的 h 指数,由此可看出 P 、 Pt_1 、 Pt_2 指数对引文、引文率的变化敏感度较高。

表7 发文量、引文量、引文率与其他文献计量指标的相关分析

Pearson	N	C	C/N	CW ₁	CW ₂	CW ₁ /N	CW ₂ /N	h	P	Pt ₁	Pt ₂
发文量 Correlation	1	0.655 **	-0.085	0.672 **	0.604 **	-0.005	-0.106	0.558 **	0.308 *	0.369 **	0.256 *
Sig. (2-tailed)		0.000	0.400	0.000	0.000	0.961	0.296	0.000	0.002	0.000	0.010
N	100	100	100	100	100	100	100	100	100	100	100
引文量 Correlation	0.655 **	1	0.584 **	0.983 **	0.978 **	0.660 **	0.514 **	0.917 **	0.886 **	0.890 **	0.838 **
Sig. (2-tailed)	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
N	100	100	100	100	100	100	100	100	100	100	100
引文率 Correlation	-0.085	0.584 **	1	0.531 **	0.612 **	0.948 **	0.959 **	0.564 **	0.865 **	0.775 **	0.872 **
Sig. (2-tailed)	0.400	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
N	100	100	100	100	100	100	100	100	100	100	100

注: **表示相关性在0.01水平上具有显著性(双侧检验), *表示相关性在0.05水平上具有显著性(双侧检验)

表8为 Pt_1 加权引文量、 Pt_1 加权引文率、 Pt_1 指数与其他计量指标的相关分析结果。从结果

数据可以看出, Pt_1 指数与除发文量之外的其他文献计量指标之间均具有较强相关性。

表8 Pt_1 指数与其他文献计量指标的相关分析

Pearson	N	C	C/N	CW ₁	CW ₁ /N	h	P	Pt ₁	Pt ₂
Pt_1 加权引文量 Correlation	0.672 **	0.983 **	0.531 **	1	0.655 **	0.925 **	0.854 **	0.905 **	0.790 **
Sig. (2-tailed)	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000
N	100	100	100	100	100	100	100	100	100
Pt_1 加权引文率 Correlation	-0.005	0.660 **	0.948 **	0.655 **	1	0.671 **	0.904 **	0.890 **	0.871 **
Sig. (2-tailed)	0.961	0.000	0.000	0.000		0.000	0.000	0.000	0.000
N	100	100	100	100	100	100	100	100	100
Pt_1 指数 Correlation	0.369 **	0.890 **	0.775 **	0.905 **	0.890 **	0.895 **	0.966 **	1	0.902 **
Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000
N	100	100	100	100	100	100	100	100	100

注: **表示相关性在0.01水平上具有显著性(双侧检验)

表 9 为 Pt_2 加权引文量、 Pt_2 加权引文率、 Pt_2 指数与其他计量指标的相关分析结果。从结果数据可以看出, Pt_2 指数与除发文量之外的其他文献计量指标之间均具有较强相关性。

表 9 Pt_2 指数与其他文献计量指标的相关分析

Pearson	N	C	C/N	CW ₂	CW ₂ /N	h	P	Pt ₁	Pt ₂
Pt_2 加权引文量 Correlation	0.604 **	0.978 **	0.612 **	1	0.592 **	0.875 **	0.882 **	0.866 **	0.887 **
Sig. (2-tailed)	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000
N	100	100	100	100	100	100	100	100	100
Pt_2 加权引文率 Correlation	-0.106	0.514 **	0.959 **	0.592 **	1	0.459 **	0.786 **	0.671 **	0.863 **
Sig. (2-tailed)	0.296	0.000	0.000	0.000		0.000	0.000	0.000	0.000
N	100	100	100	100	100	100	100	100	100
Pt_2 指数 Correlation	0.256 *	0.838 **	0.872 **	0.887 **	0.863 **	0.786 **	0.965 **	0.902 **	1
Sig. (2-tailed)	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
N	100	100	100	100	100	100	100	100	100

注: ** 表示相关性在 0.01 水平上具有显著性(双侧检验), * 表示相关性在 0.05 水平上具有显著性(双侧检验)

相关分析可以看出, Pt_1 指数、 Pt_2 指数的评价结果与 h 指数和 P 指数评价结果皆具有较高相关性, 与 P 指数的相关性尤高。

4 结论与讨论

本文在 P 指数研究成果的基础上, 考虑时间的稀释效应与凝聚效应, 分别构建了 Pt_1 指数和 Pt_2 指数, 选择图书情报与文献学领域 h 指数排名前 100 位的学者为样本, 以 CSSCI 收录的文献为对象, 进行实证分析, 得出以下研究结论。

(1) 与 h 指数相比, 两种 Pt 指数更具区分度和灵敏度。h 指数进行科研评价最明显的一个缺陷就是缺乏足够的区分度。从洛特卡定律可知, 在科学研究中高产作者较少, 中产和低产作者却非常多, 而这些中产和低产作者中会有许多作者具有相同的 h 指数, 此时 h 指数就无法对这部分作者进行合理的评价。Pt 指数基于引文量和引文率构建的经验公式, 其计算结果通常是小数, 能对具有相同 h 指数的作者进行适当的区分评价, 且任何新的引用都会对实际评价产生影响, 体现出较高的灵敏性。

(2) 与 P 指数相比, Pt_1 指数更能识别发文时间较短的潜在高价值文献。在 P 指数的评价中, 发文一年被引 10 次和发文五年被引 10 次的文献影响力是相同的, 但事实上两者的学术影响力是不相同的, 同时发文一年被引 10 次的文献在第二年可能会得到更多被引, 是潜在的高价值文献。通过 Pt_1 指数反映时间的稀释效应, 不仅能区分相同引文量的文献, 而且更能识别出发文时间较短的潜在高价值文献。

(3) 与 P 指数相比, Pt_2 指数更能识别高影响力的经典文献。被引量同为 50 次的两篇文献, 在 P 指数的评价中这两篇文献的影响力是相同的, 但在 Pt_2 指数的评价中, 半衰期外的文献会被赋予更高权重, 这就使得文献年龄大于半衰期的文献具有更高的加权引文量, 从而识别出这类高价值的经典文献。

(4) 对于半衰期较长的学科领域, 作为半衰期外引文调节因子的权重通常也不会很高, 这就使得学科领域中文献 Pt_2 加权被引量与原被引量相差不大, 也使得 Pt_2 指数的评价结果与 P 指数相关性很高。因此, 如果用 Pt_2 指数来评价学科领域半衰期较长的科研人员、学科期刊或

者科研组织时,不如P指数方便快捷。

(5)从相关分析结果来看, P_{t_1} 和 P_{t_2} 指数与发文量的相关性都较小,而与引文量的相关性较大,这可能会给评价带来负面影响,促使部分人员、组织仅关注引文量而忽视发文量,因而在实际中需要与其他一些指标综合起来进行分析评价。

本文仅以图书情报与文献学领域为例,对

构建的 P_{t_1} 和 P_{t_2} 指数进行适用性验证,在其他学科领域是否也适用,需要进一步验证。另外, P_{t_2} 指数以半衰期作为阈值划分标准是否合理,是否还有其他更好的划分方法,对评价对象中的低被引或零被引的文献是否需要作适当的处理,以及如何处理,等等,这些问题在今后的研究中需要深入探讨。

参考文献

- [1] 叶鹰.国际学术评价指标研究现状及发展综述[J].情报学报,2014,33(2):215-224.(Ye Ying.The research progress and developing perspective of assessment indicators[J].Journal of the China Society for Scientific and Technical Information,2014,33(2):215-224.)
- [2] Lee C J,Sugimoto C R,Zhang G,et al. Bias in peer review[J]. Journal of the American Society for Information Science and Technology,2013,64(1):2-17.
- [3] 韩毅,张克菊,金碧辉.引文网络分析的方法整合研究进展[J].中国图书馆学报,2010,36(4):83-89.(Han Yi,Zhang Keju,Jin Bihui. Methodology integration of citation network analysis[J]. Journal of Library Science in China,2010,36(4):83-89.)
- [4] Hirsch J E. An index to quantify an individual's scientific research output [J].Proceedings of the National Academy of Sciences of the United States of America,2005,102(46):16569-16572.
- [5] Prathap G. A performance index approach to library collection[J]. Performance Measurement and Metrics,2010,11(3):259-265.
- [6] 王志军,郑德俊.P指数运用于人才评价的有效性实证研究[J].图书情报工作,2012,56(14):93-97.(Wang Zhijun,Zheng Dejun.Application of P-index to evaluating individual's performances[J]. Library and Information Service,2012,56(14):93-97.)
- [7] Schubert A,Glänzel W. A systematic analysis of Hirsch-type indices for journal[J]. Journal of Informetrics,2007,1(3):179-184.
- [8] Glänzel W. On some new bibliometric applications of statistics related to the h-index [J]. Scientometrics,2008,77(1):187-196.
- [9] Csajbok E,Berhidi A,Vasas L,et al. Hirsch-index for countries based on essential science indicators data[J]. Scientometrics,2007,73(1):91-117.
- [10] Prathap G. Is there a place for a mock h-index?[J]. Scientometrics,2010,84(1):153-165.
- [11] Prathap G. The 100 most prolific economists using the p-index[J]. Scientometrics,2010,84(1):167-172.
- [12] 夏慧,韩毅.一个新的综合性科技评价指标——p指数研究综述[J].图书情报工作,2014,58(8):128-132.(Xia Hui,Han Yi.A new comprehensive metric indicator to science and technology:a review on p-index[J]. Library and Information Service,2014,58(8):128-132)
- [13] Prathap G. The iCE approach for journal evaluation [J]. Scientometrics,2010,85(2):561-565.
- [14] Prathap G. Qualifying scholarly impact using an iCX (impact-Citations-Exergy) analysis [J]. Journal of Library & Information Technology,2011,31(5):382-386.
- [15] Prathap G. The fractional and harmonic p-indices for multiple authorship [J]. Scientometrics,2011,86(2):

- 239-244.
- [16] Prathap G. A Thermodynamic explanation for the Glänzel-Schubert model for the h-Index[J]. Journal of the American Society for Information Science and Technology, 2011, 62(5): 992-994.
- [17] Prathap G. The energy-exergy-entropy (or EEE) sequences in bibliometric assessment[J]. Scientometrics, 2011, 87(3): 515-524.
- [18] Prathap G. A three-class, three-dimensional bibliometric performance indicator[J]. Journal of the Association for Information Science and Technology, 2014, 65(7): 1506-1508.
- [19] 张正慧, 郑德俊. P 指数用于学术期刊评价的适用性分析[J]. 图书情报工作, 2013, 57(19): 121-126. (Zhang Zhenghui, Zheng Dejun. The validity of P-index in evaluating Chinese scholarly journals[J]. Library and Information Service, 2013, 57(19): 121-126.)
- [20] 王圣洁, 蒋旭, 何晓庆. 期刊 P 指数与其他评价指标相关性探讨[J]. 中国科技期刊研究, 2014, 25(11): 1404-1407. (Wang Shengjie, Jiang Xu, He Xiaoqing. Relation discussion between the P-index and other metric index about journals[J]. Chinese Journal of Scientific and Technical Periodicals, 2014, 25(11): 1404-1407.)
- [21] Petersen A M, Succi S. The Z-index: a geometric representation of productivity and impact which accounts for information in the entire rank-citation profile[J]. Journal of Informetrics, 2013, 7(4): 823-832.
- [22] 朱惠, 邓三鸿, 杨建林. 我国图书情报学领域博士生导师 H 指数分析[J]. 图书与情报, 2013(5): 57-66. (Zhu Hui, Deng Sanhong, Yang Jianlin. H-index analysis of PhD supervisors in Library and Information Science fields[J]. Library & Information, 2013(5): 57-66.)
- [23] 邱均平, 王菲菲. 基于引证关系的国内情报学领域作者研究活力与影响力分析[J]. 图书馆论坛, 2011, 31(6): 51-60. (Qiu Junping, Wang Feifei. Analysis about research activity and influence of authors in China based on citation relationship[J]. Library Tribune, 2011, 31(6): 51-60.)
- [24] 邱均平, 周春雷. 发文量和 h 指数结合的高影响力作者评选方法研究——以图书情报学为例的实证分析[J]. 图书馆论坛, 2007, 28(6): 44-49. (Qiu Junping, Zhou Chunlei. The method of selecting high influence authors using published and h-index—a case study of the library and information field[J]. Library Tribune, 2007, 28(6): 44-49.)
- [25] 丁楠. 基于 CSSCI 的文科学者 h 指数实证研究[D]. 杭州: 浙江大学, 2007. (Ding Nan. A CSSCI-based positive study on the h-index of Humanities and Social Sciences scholars[D]. Hangzhou: Zhejiang University, 2007.)

韩毅 西南大学计算机与信息科学学院教授, 硕士生导师。重庆 400715。

夏慧 西南大学计算机与信息科学学院情报学硕士研究生。重庆 400715。

(收稿日期: 2015-06-28; 修回日期: 2015-07-24)