

我国省级公共图书馆科研产出的 h 指数测度

赵 星 黄 娟 彭晓东

摘 要 对我国 31 个省级公共图书馆 1999 – 2008 年的科研产出状况及 h 指数特性进行了实证研究,结果表明:上海图书馆和广东省立中山图书馆具有较突出的论文产出能力和学术影响力,河南省图书馆进步较大;各馆馆员论文 h 指数总体呈正态分布,东西部差距大,东中部差异小,不同时段论文产出规模和科研能力分级态势稳定;与高校图书馆相比,省级公共图书馆馆员论文 h 指数平均值虽低,但峰值高,各馆研究模式呈现“百花齐放”格局;h 指数对图书馆科研产出的测评具有代表性和综合性,可激励图书馆员从事创新性研究。图 5。表 2。参考文献 20。

关键词 省级公共图书馆 科研能力评价 h 指数 科学计量 信息计量

分类号 G350

ABSTRACT We conducted an empirical study of the research output and h-index features of 31 provincial public libraries in China during 1996 – 2000. The result shows that the Shanghai Library and the Sun Yat-sen Library of Guangdong Province have outstanding research paper output capacity and academic impacts, and the Henan Provincial Public Library has made an obvious progress. On the whole, the h-index of provincial public libraries presents a normal distribution, with considerable difference between the east and west regions, but a smaller gap between the east and middle regions. In different stages, the general pattern of paper output and research capacity is maintained. Compared with academic libraries, the h-index of provincial public libraries has lower mean value but higher peak value, with various libraries presenting a large variety of research models. The h-index measure for research output of libraries is a representative and comprehensive method, which may encourage librarians to engage in creative researches. 5 figs. 2 tabs. 20 refs.

KEY WORDS Provincial public library. Research evaluation. H-index. Scientometrics. Informetrics.

CLASS NUMBER G350

1 引言

公共图书馆是图书馆学术研究的重要力量,学术研究也日益成为公共图书馆事业发展的重要推动力^[1-2]。宋萍^[3]曾用 1998 – 2002 年在图书馆学中文核心期刊上发表论文数为指标,研究了省级公共图书馆的科研实力。但目前对于我国公共图书馆科研工作的研究,以定性探讨为主,定量实证较少,以论文数量作为科研评价指标的合理性近年来备受争议,公共图

书馆科研产出的测度应寻求更有效方法。笔者最近研究发现,h 指数可更好地测评图书馆的科研能力^①。此方法一定程度上弥补了论文数量和被引频次等指标的不足,能综合衡量论文的数量和影响力。

本文用 h 指数研究我国省级公共图书馆的科研产出状况。我们首先介绍图书馆馆员论文的 h 指数计算方法,并计算 1999 – 2008 年我国 31 个省级公共图书馆馆员论文的 h 指数及其它计量指标,然后分别讨论我国省级公共图书馆馆员论文 h 指数的总体情况、各指标间联系及

① 该结论来自笔者另一针对高校图书馆科研能力的 h 指数测度研究,已撰成论文《高校图书馆科研能力的 h 指数测度》即将在《图书馆杂志》刊出。本文引用了其中的一些研究数据和结论。

与高校图书馆馆员论文的 h 指数差异,最后总结 10 年间我国省级公共图书馆科研产出状况及 h 指数特性。

2 方法与数据

Hirsch 提出 h 指数的原始论文^[4]已在 Web of Science 中被引用了 264 次(查询时间为 2009 年 4 月 1 日)。由于具有诸多优点^[5],h 指数已被用于科学家、期刊^[6-7]、研究团队^[8]、科研机

构^[9]、专利^[10]和科学基金^[11]等不同领域的评价,成为国际信息计量学近年来主要的新兴热点^[12],并有望发展为下一代核心科学评价参数^[13-14]。

根据 h 指数的定义^[4],图书馆馆员论文 h 指数可描述为:某一时段内,某图书馆所属馆员发表的论文集合中,至少有 h 篇论文的被引次数不少于 h 次。以重庆图书馆为例的 h 指数计算方法如表 1 所示。

表 1 重庆图书馆论文 h 指数计算方法

序号	论文题目及作者	论文被引频次	判定
1	数字图书馆与人才问题(邵康庆)	18	1 < 18
2	发展数字化公共图书馆的著作权问题研究(张颖,张力)	8	2 < 8
3	从学科交叉看情报学发展(曾妍)	8	3 < 8
4	对网络环境下图书采访工作的思考(严国秀)	6	4 < 6
5	图书馆未来发展及功能的变迁(林滨)	5	5 = 5, h 指数为 5
6	论图书馆学与信息学教育的职业取向(高英兰,熊忠华)	4	6 > 4
7	公共图书馆特色服务的可持续发展(邵康庆,王祝康)	4	7 > 4
...	...	...	...

注:此表中数据的收集和计算过程为:进入 CNKI 引文数据库,检索项选择“被引文献作者机构”,关键词为图书馆名(此处应用逻辑“或”包含“重庆图书馆”和“重庆市图书馆”两个检索关键词),范围勾选全部学科,时段选择 1999 - 2008 年,匹配模式为“模糊”,排序按“被引频次”,检索得出此时段重庆图书馆的作者发表的有被引用过的论文及其被引频次,从而可找出序号 h 值,使得前 h 篇论文被引频次都大于或等于 h,而序号为(h + 1)的论文被引频次小于(h + 1)。使用其它引文数据库进行计算也可按此方法灵活运用。

表 1 中,对重庆图书馆馆员论文 h 指数有直接贡献的是被引频次最高的 5 篇论文,图书馆馆员论文的 h 指数实际上是由各馆作者发表论文中影响力最大的一批作品所决定。可见,要提高图书馆馆员论文的 h 指数,单是论文数量的累积作用不大。以 1999 年之后重庆图书馆为例,2008 年后该馆新发表论文的被引频次若等于或少于 5,无论数量有多少,都无法改变该馆

1999 年后的 h 指数,而只有在产出了被引频次高于 5 的论文时,才有可能提升该馆 h 指数^①。

利用 CNKI 引文数据库和表 1 方法,我们分别计算了 1999 - 2003 年、2004 - 2008 年和 1999 - 2008 年三个时段我国 31 个省级公共图书馆(不含港澳台地区)馆员论文的 h 指数,并统计了各馆馆员论文的总被引频次。再利用 CNKI 全文数据库获取在此时段内各馆馆员发表论文

① 仅是“有可能”是因为 1999 - 2008 年期间该馆作者发表论文的被引频次也可能随时间推移而增长,并使原来的 h 指数提升,若新发表论文的被引频次不能超过提升后的 h 指数,也不能对该馆的 h 指数增加产生直接贡献。

数量,并用总被引频次除以论文数量得到篇均被引频次。检索时间为 2008 年 12 月,论文和引文截止时间均为 2008 年 12 月。检索时尽可能考虑了统计时段内馆名的更改或多种称谓,如广东省立中山图书馆有曾用名“广东省中山图书馆”;又如检索浙江图书馆的数据时,还使用了“浙江省图书馆”为关键词。结果数据包括了 31 个省级公共图书馆在 10 年间发表的 5756 篇论文及 16624 条被引记录。另外,我们也选取了 h 指数前 31 位的高校图书馆为对象,按本研究的统计时段和方法重新计算了各馆馆员论文的 h 指数,以便进行省级公共图书馆与高校图书馆馆员论文 h 指数的实证比较。

3 分析与讨论

3.1 省级公共图书馆馆员论文 h 指数的总体情况

本文将 1999 - 2008 年省级公共图书馆馆员的论文数量、总被引频次、篇均被引和 h 指数分别以 p、c、ac 和 h 表示,1999 - 2003 年(以下简称时段 1)的各指标分别用 p1、c1、ac1 和 h1 表示,2004 - 2008 年(以下简称时段 2)的各指标为 p2、c2、ac2 和 h2,三个时段论文各指标的代表性结果见表 2。

表 2 1999 - 2008 年 h 指数前 20 位的省级公共图书馆馆员论文相关数据(按 h 指数降序排列)

馆 名	1999 - 2008 年				1999 - 2003 年(时段 1)				2004 - 2008 年(时段 2)			
	p	c	ac	h	p1	c1	ac1	h1	p2	c2	ac2	h2
上海图书馆	631	3582	5.68	30	327	2742	8.39	28	304	840	2.76	13
广东省立中山图书馆	372	2174	5.84	23	136	1505	11.07	22	236	669	2.83	12
湖南图书馆	278	1263	4.54	18	144	928	6.44	14	134	335	2.50	9
天津图书馆	329	986	3.00	15	195	715	3.67	14	134	271	2.02	9
浙江图书馆	299	730	2.44	14	149	576	3.87	13	150	154	1.03	6
辽宁省图书馆	360	1002	2.78	14	183	752	4.11	13	177	250	1.41	8
黑龙江省图书馆	358	804	2.25	13	178	672	3.78	12	180	132	0.73	6
河南省图书馆	342	649	1.90	12	158	399	2.53	10	184	250	1.36	9
南京图书馆	434	946	2.18	11	203	704	3.47	11	231	242	1.05	7
福建省图书馆	153	502	3.28	11	63	362	5.75	10	90	140	1.56	6
吉林省图书馆	246	461	1.87	10	111	328	2.95	10	135	133	0.99	5
陕西省图书馆	129	279	2.16	9	50	228	4.56	9	79	51	0.65	4
首都图书馆	81	207	2.56	9	52	175	3.37	9	29	32	1.10	3
山西省图书馆	216	430	1.99	9	132	363	2.75	9	84	67	0.80	4
甘肃省图书馆	175	397	2.27	9	95	314	3.31	9	80	83	1.04	5
山东省图书馆	167	252	1.51	8	101	203	2.01	8	66	49	0.74	3
江西省图书馆	181	356	1.97	8	107	257	2.40	8	74	99	1.34	5
安徽省图书馆	147	281	1.91	8	68	177	2.60	7	79	104	1.32	5
湖北省图书馆	76	267	3.51	7	38	211	5.55	7	38	56	1.47	4
贵州省图书馆	114	193	1.69	7	30	132	4.40	5	84	61	0.73	3

1999 - 2008 年 10 年间,我国省级公共图书馆平均每馆发表被 CNKI 收录的论文 185.68 篇,平均被引频次为 536.26 次,平均篇均被引

2.21 次,平均 h 指数为 9.16,大部分省级公共图书馆的科研参与度较高。由表 2 可见,上海图书馆和广东省立中山图书馆等馆具有较高的馆员

论文 h 指数。其中上海图书馆 1999–2008 年馆员论文 h 指数为 30, 表明该馆在此期间共有 30 篇论文被引次数不少于 30 次。另外, 该馆的论文数量和被引频次也具明显优势, 显示出较强的科研规模和实力。从数值分布上分析, Kolmogorov-Smirnov 检验表明省级公共图书馆的馆员论文 h 指数总体上服从正态分布 (双尾显著性概率为 $0.44 > 0.05$), 分布直方图见图 1。由图 1 看出, 各馆馆员论文的 h 指数较集中地分布于 0–19 之间, 馆员论文 h 指数为 8 和 9 的图书馆较多, 高值较少。

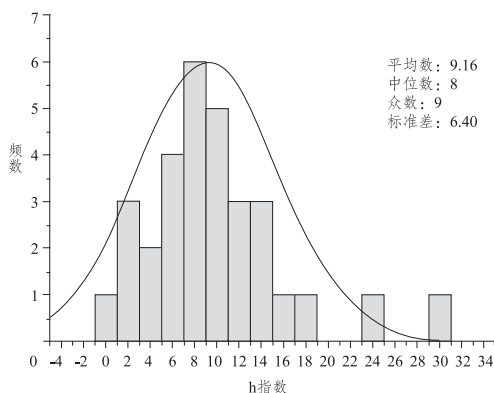


图 1 1999–2008 年省级公共图书馆馆员论文 h 指数描述

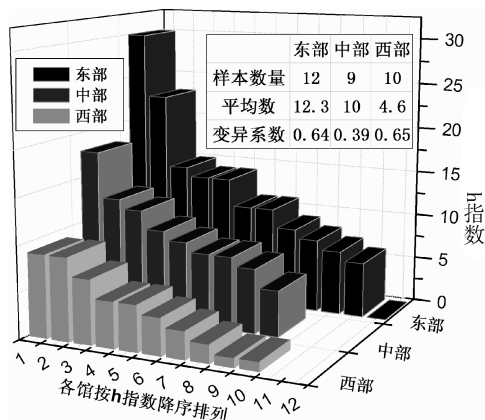


图 2 1999–2008 年东、中、西部省级公共图书馆的馆员论文 h 指数比较

图 2 的地域分析显示, 从东部到西部, h 指数均值呈下降趋势, 东部和中部差距较小, 西部与其他两个区域差距较大。就区域内部而言, 东部和西部各馆馆员论文的 h 指数差别较大, 而中部各馆差异相对较小。尽管东部高馆员论文 h 指数的图书馆较多, 但中部和西部也有一些图书馆在馆员论文 h 指数上表现良好, 如中部的湖南图书馆馆员论文 h 指数位列所有省级公共图书馆第三, 西部的陕西省图书馆和甘肃省图书馆馆员论文 h 指数也位列所有省级公共图书馆的前半区。

3.2 省级公共图书馆馆员论文 h 指数与其他指标的联系

论文数量 p 、被引频次 c 和篇均被引 ac 都是常用的文献计量指标, 分别代表论文集合的规模、影响力和篇均影响力, 而 h 指数则由论文集合中高影响力论文所决定。图 3 表明, h 指数与其他指标的正相关关系较明显, 代表各指标的能力相对较强。这揭示了一个现象: 高被引论文多 (即馆员论文 h 指数高) 的省级公共图书馆, 通常发表的论文较多, 而且这些论文的总体和平均影响力都更强。

经过不同模型的测试和比较, 省级公共图书馆馆员论文 h 指数与各指标的拟合关系式为:

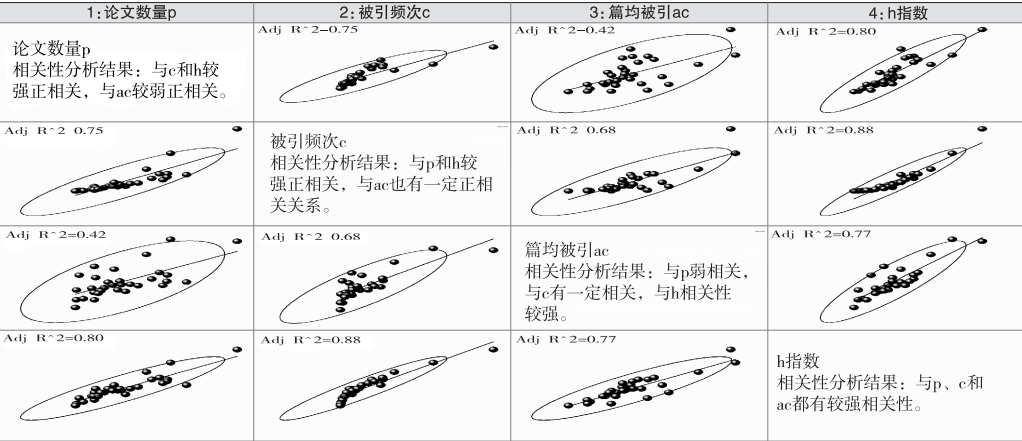
$$h = 0.16p^{0.78} (R^2 = 0.80, P < 0.01) \quad (1)$$

$$h = 0.42c^{0.52} (R^2 = 0.97, P < 0.01) \quad (2)$$

$$h = 4.18ac (R^2 = 0.77, P < 0.01) \quad (3)$$

公式 (1)–(3) 说明省级公共图书馆馆员论文 h 指数与各指标间存在拟合优度较高的经验关系, 其与论文数量和被引频次为幂律关系, 而与篇均被引近似直线关系。图 3 和公式 (2) 都显示 h 指数与被引频次关系密切, 且省级公共图书馆的总体影响力与高影响力论文的数量有较强的相关性。这一方面是因为高影响力论文的被引频次是各馆馆员论文总被引频次的重要组成部分, 如上海图书馆和广东省立中山图书馆作者发表的对于 h 指数有直接贡献的 30 篇和 23 篇论文中, 数量只分别占本馆论文的 5% 和 6%, 但被引频次所占比例高达 49% 和 48%。另一方面, 科研实力强的省级公共图书馆影响力

与本馆作者的影响力之间存在相互促进的机制,会使得该馆作者的研究受到更多关注并获
得更大影响力,进而产生更多的高被引论文。



注: (1) Adj R² 为修正相关系数; (2) 图中椭圆表示置信范围; (3) 处于各子图最右上端的两个数据点为上
海图书馆和广东省立中山图书馆。

图 3 1999 - 2008 年省级公共图书馆馆员论文各指标相关程度分析

图 3 中有一个有趣的现象, 尽管上海图书馆和广东省立中山图书馆的数据点并未过于偏离拟合曲线, 但上海图书馆的数据点却常在置信椭圆之外, 广东省立中山图书馆的数据点也常在置信椭圆边界附近。这说明两个图书馆的指标虽然基本符合整体规律, 但数值上与数据集主体有较大差别, 再次体现了两个馆在科研产出规模和影响力上的突出优势。

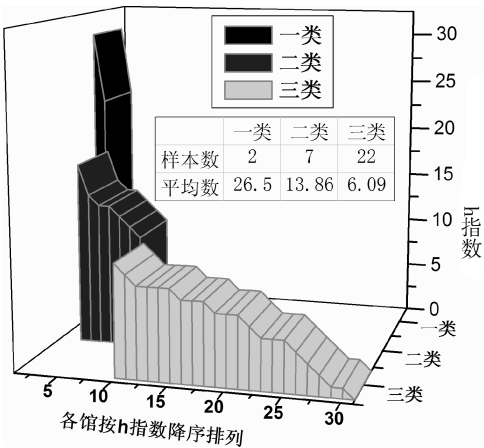


图 4 指标类平均法聚类与 h 指数排序结果比较

依据各馆馆员论文的 p、c、ac 和 h 指数四种指标数据, 采用类平均法进行聚类分析, 结果 31 个省级公共馆按指标表现被分为三类, 其中上海图书馆和广东省立中山图书馆为第一类; 湖南图书馆和天津图书馆等 7 馆为第二类; 其余 22 馆为第三类。将聚类结果与 h 指数的排序进行比较 (见图 4), 发现四种指标聚类的分类结果完全符合 h 指数的排序, 即 h 指数前两位的图书馆恰好是第一类, 3—9 位的图书馆恰好是第二类, 其余为第三类。而三种其它指标的排序与聚类结果却有一定差异。此结果说明, h 指数的测评具有代表性, 能较好地反映四种指标的综合结果。

为比较各馆 1999 - 2003 年和 2004 - 2008 年两个时段的指标综合变化, 我们采用类平均法对 31 个省级公共馆的两个时段数据分别进行聚类分析。结果发现各馆在两个时段也都被分为三类, 省级公共图书馆的总体科研能力分级态势较稳定。其中, 上海图书馆和广东省立中山图书馆在两个时段都共同构成第一类; 时段 1 的第二类包含 6 个图书馆; 而时段 2 包含 5 个图书馆。湖南图书馆、南京图书馆、天津图书馆和辽宁省图书馆在两个时段都属于第二类, 学术

论文成果丰富且影响力较大。河南省图书馆在 10 年间取得了较大进步,从时段 1 的第三类跃居时段 2 的第二类。

3.3 省级公共图书馆与高校图书馆的馆员论文 h 指数比较

高校图书馆也是图书馆学研究的重要力量,我们的前期研究曾探讨了高校图书馆馆员论文 h 指数的部分特点,本小节将结合两个研究的数据比较省级公共图书馆馆员论文 h 指数与高校图书馆馆员论文 h 指数的差异。

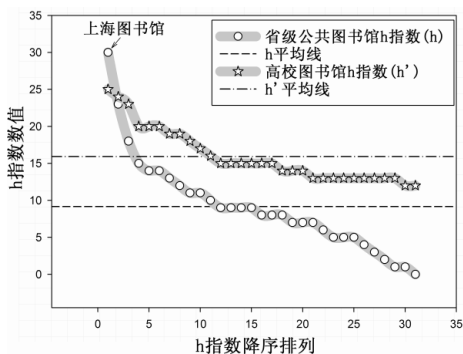


图 5 1999-2008 年省级公共图书馆与部分高校图书馆的馆员论文 h 指数比较

由于本文数据只包含 31 个省级公共图书馆,而高校图书馆较多,故我们选取的是前期研究中馆员论文 h 指数最高的 31 个高校图书馆作为样本,并按本文方法和统计时段重新计算各馆馆员论文的 h 指数。图 5 是 1999-2008 年省级公共图书馆与前述样本高校图书馆的馆员论文 h 指数比较。可见,作为公共图书馆代表的 31 个省级公共图书馆馆员论文 h 指数平均水平要低于 31 个高校图书馆,有 3 个省级公共图书馆的馆员论文 h 指数超过样本高校图书馆的平均线。若将两组图书馆馆员论文 h 指数合并再降序排列,前 10 位中只有两个图书馆属于公共图书馆系统。这可能是由于高校图书馆特别是部分名校的图书馆在研究上得到了所在学校各方面的支撑或合作,而公共图书馆这方面的资源则相对不足。另外,部分经济欠发达地区由于经费和人力资源稀缺等原因,科研投入较少

或激励机制缺失,都导致 h 指数较低,因此整体平均数不如高校图书馆的平均值。当然,在 h 指数峰值上,上海图书馆作为馆员论文 h 指数最高的省级公共图书馆,h 指数也同时明显高于各高校图书馆,体现了出众的高水平学术论文产出能力和创新能力。

之前的研究显示,高校图书馆馆员论文 h 指数较高的原因主要有两种:科研队伍整体水平高或少数学术带头人的贡献突出。但分析了本文中馆员论文 h 指数排名前 10 位的公共图书馆作者后发现,与高校图书馆不同,馆员论文 h 指数较高的省级公共图书馆的学术研究更为“百花齐放”,如上海图书馆和广东省立中山图书馆都有多位学者产出了多篇高影响力论文,进而对本馆的 h 指数作出贡献,显示了活跃的群体科研风气。而其它排名靠前的图书馆也与之相似,对本馆 h 指数作出贡献的论文分别属于多位作者,而并不集中于少数两三人。

如公式(4)-(6)所示,高校图书馆的馆员论文数量 p、被引频次 c 和篇均被引 ac 与 h 指数的拟合规律为幂律关系。

$$h = 0.61p^{0.55} (R^2 = 0.69, P < 0.01) \quad (4)$$

$$h = 0.60c^{0.47} (R^2 = 0.96, P < 0.01) \quad (5)$$

$$h = 4.18ac^{1.15} (R^2 = 0.35, P < 0.01) \quad (6)$$

可见,省级公共图书馆馆员论文的 p 和 c 与 h 指数的拟合规律与高校图书馆相似,但公式(3)显示,省级公共图书馆馆员论文的 ac 与 h 指数的经验关系更接近直线而非公式(6)的幂律,这说明省级公共图书馆馆员论文篇均被引的增量与 h 指数增量关系较稳定,粗略而言,篇均被引增加 1 的同时 h 指数也会增加 5。与高校图书馆的数据拟合规律相似,公式(1)和公式(2)的幂指数都在 0~1 之间,表明省级公共图书馆馆员论文数量和总被引的增长对于 h 指数增长的贡献也存在“规模效应递减”现象。这说明至少在图书馆馆员论文 h 指数上,“规模效应递减”现象有一定普适性。前面已论证了单是论文数量的累积对各馆馆员论文 h 指数的提升作用不大,结合此现象则可说明,即使被引频次大量增加,但若增量主要来自新发表的低被引论文的累积,对 h 指数的提升意义也不大。因

此,要提升图书馆馆员论文的h指数,更重要的是要进行超越本馆现有水平的更有影响力的研究。需要说明的是,使用h指数为评价方法,目的是激励图书馆加大原创性研究的投入力度,而不是刻意追求论文的产出。

4 结语

本文以图书馆h指数为主要方法,研究分析了1999-2008年我国31个省级公共图书馆的科研状况及h指数特性。当然,h指数仅是科研绩效中论文方面的评价参数,论文主要体现基础研究和基础应用研究成果,更全面的科研能力测度还需考虑著作、专利、基金课题和社会影响等因素,实践中应注意将h指数与其它方法结合使用。下一步可将研究范围扩展到其它级别公共图书馆或专业图书馆。世界范围内各国公共图书馆科研态势的分布情况,也可通过类似本文的h指数方法和Web of Science或Scopus等数据库得出。另外,国内外学者在h指数基础上提出了一系列新方法,如Egghe的g指数^[15],Kosmulski的h(2)指数^[16],金碧辉和梁立明等人的A、R与AR指数^[17],叶鹰的F指数^[18]以及Rousseau和叶鹰的动态h指数^[19]等,这些新指数各具特点,也可尝试将之用于图书馆科研产出状况的测度。

参考文献:

- [1] 邹琦. 公共图书馆学术研究应与实际工作互动[J]. 图书馆, 2008(3): 90-91.
- [2] 钟海珍. 公共图书馆学术研究现状之忧与思[J]. 图书馆论坛, 2006, 26(5): 68-70.
- [3] 宋萍. 省级公共图书馆专业研究实力分析[J]. 图书与情报, 2003, 4: 58-61.
- [4] Hirsch J. An index to quantify an individual's scientific research output [J]. PNAS, 2005, 102(46): 16569-16572.
- [5] Bornmann L, Daniel HD. What do we know about the h-index? [J]. JASIST, 2007, 58: 1381-1385.
- [6] Braun T, Glänzel W, Schubert A. A Hirsch-type index for journals [J]. Scientometrics, 2006, 69(1): 169-173.
- [7] 赵基明, 舒明全. 基于CSSCI的《中国图书馆学报》h指数及核心作者测定[J]. 中国图书馆学报, 2008, 34(174): 98-102.
- [8] Van Raan A. Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups [J]. Scientometrics, 2006, 67(3): 491-502.
- [9] Arencibia-Jorge R, Barrios-Almaguer I, Fernandez-Hernandez S, et al. Applying successive H indices in the institutional evaluation: A case study [J]. JASIST, 2008, 59(1): 155-157.
- [10] Guan JC, Gao X. Exploring the h-index at patent level [J]. JASIST, 2009, 60(1): 35-40.
- [11] 赵星, 高小强, 何培. 科学基金h指数: 基金论文成果数量与影响力的综合衡量[J]. 中国科学基金, 2009(1): 15-19.
- [12] 赵星, 高小强, 郭吉安, 陶乃航. 基于主题词频和g指数的研究热点分析方法[J]. 图书情报工作, 2009, 53(2): 31-34.
- [13] 叶鹰. h指数和类h指数的机理分析与实证研究导引[J]. 大学图书馆学报, 2007(5): 2-5.
- [14] 叶鹰. 图书情报学前沿研究领域选评[J]. 中国图书馆学报, 2008, 34(176): 63-70.
- [15] Egghe L. Theory and practise of the g-index [J]. Scientometrics, 2006, 69(1): 131-152.
- [16] Kosmulski M. A new Hirsch-type index saves time and works equally well as the original h-index [J]. ISSI Newsletter, 2006, 2(3): 4-6.
- [17] Jin B, Liang L, Rousseau R. The R- and AR indices: Complementing the h-index [J]. Chinese Science Bulletin, 2007, 52(6): 855-863.
- [18] Ye F Y. F-index: a new index for academic ranking at all-round levels. Analysis and Creation, China Society Press, 2008 (2nd ed.): 101-120.
- [19] Rousseau R, Ye F Y. A proposal for a dynamic h-type index [J]. JASIST, 2008, 59(11): 1853-1855.

赵星 黄娟 重庆大学经济与工商管理学院情报学硕士研究生。通讯地址: 重庆大学。邮编 430030。

彭晓东 重庆大学图书馆馆长, 教授, 博士生导师。通讯地址同上。

(收稿日期: 2009-03-12; 修回日期: 2009-04-01)