

F 指数:一种改进的学术影响力测量方法

王传毅 俞寅威

(清华大学教育研究院,北京,100084)

[摘要] 本研究提出了一种改进的学术影响力测量方法——F 指数,并以哲学、法学、统计与概率、机械工程、作物学、传染病学以及交叉学科领域的学者为样本,实证检验了 F 指数对学术影响力的评价效果。研究发现,F 指数在跨学科领域的评价中,能够调节不同领域学术影响的差异程度,较 h 指数、p 指数、FWCI 具有更佳的分度;在单一学科领域评价中,与 h 指数、p 指数、FWCI 具有高度一致性,也具有更佳的分度,是对现有指标的进一步改进。F 指数可用于学者、科研团队及学术组织的学术影响力评价。

[关键词] F 指数 学术影响力 学术评价 跨学科评价 评价成本 评价精度

[中图分类号] G350 [文献标识码] A [文章编号] 2095-2171(2022)04-0024-09

DOI: 10.13365/j.jirm.2022.04.024

F-index: An Improved Method of Measuring Academic Influence

Wang Chuanyi Yu Yinwei

(Institute of Education, Tsinghua University, Beijing 100084)

[Abstract] This research presents F index as an improved method of measuring academic influence. By using scholars in Philosophy, Laws, Statistics and Probability, Mechanical Engineering, Agronomy and Crop Science, Infectious Diseases and Multidisciplinary fields as samples, the effect of F-index on the measurement of academic influence is empirically tested. The result indicates that in the field of multidisciplinary evaluation, F-index shows a better discernment than h-index, p-index and FWCI with the capability of adjusting different standards of academic influence among different fields. In the evaluation of single subject, F-index has a high consistency with h-index, p-index and FWCI but also shows a better discernment, it is a further improvement based on the existing indicators.

[Keywords] F-index; Academic influence; Academic evaluation; Multidisciplinary evaluation; Evaluation cost; Evaluation accuracy

1 引言

学术影响力是评价科研成果、学者及学术组织的重要维度。h 指数是学术影响力常用的量化评价办法^[1]。h 指数兼具考察被引

频次和论文数量两个方面,既可以遏制片面追求论文数量的不良倾向,同时又能够激发科研人员撰写高质量论文的热情^[2]。但 h 指数也存在明显缺陷,一是区分度低,特别是对普通科学家的学术影响力进行比较时, h 指

[基金项目] 国家自然科学基金项目“基于海量数据的博士生科研网络对学术生产力影响的混合研究”(71904100)

[作者简介] 王传毅,博士,副教授,研究方向为高等教育评价,Email:wcy1985@tsinghua.edu.cn;俞寅威,硕士生。

本文引用格式:王传毅,俞寅威.F 指数:一种改进的学术影响力测量方法[J].信息资源管理学报,2022,12(3):24-32.

数并不具备鉴别力^[3]；二是信息损失，h指数不仅忽视了高被引论文的贡献，也完全忽视了h指数以下的论文数量及被引频次^[4]；三是只升不降，h指数是一个累积指标，对年轻的科研工作者不利，同时允许科学家“吃老本”，无法反映科学家研究活力的衰退情况^[5]。

为改进h指数的缺陷，国内外学者先后提出了改进型h指数，即类h(h-type)指数，主要包括h(2)指数^[6]、g指数^[7]、A指数^[8]、R指数^[9]、AR指数^[9]、 h_g 指数^[10]、 h_k 指数^[11]、A+指数^[12]等等，其中部分指标是在h指数的基础上引入新的变量(如论文总被引量、平均被引量、h核心论文被引量、h核心论文平均被引量、论文最高被引量等)，还有部分指标是对h指数进行一定的数学运算(如对论文被引数据进行平方，对几项指标之积求平方根等)，这些类h指数都不同程度地提高了h指数的区分能力，但定义较为复杂，可操作性并不强。真正理想的评价指标需要同时考虑到评价成本与评价精度，在保证必要的评价精度的前提下，指标越直观、越简单、越容易计算越好^[13]。此外，对于h指数的累积性问题，也未能完全解决。

2010年Prathap在Glanzel的研究基础上，提出了p指数^[14]，定义为： $p = \left(\frac{C^2}{N}\right)^{\frac{1}{3}}$ ，其中N为论文数量，C为论文被引次数，Prathap认为该指标与其他类h指数相比，计算过程简单，特别是当期刊引文存在高引现象(单篇文章的引文大大超过了h值)或长尾现象(大量未被引用的文章)时，能够表现出更强的敏感性。Prathap还将其与物理学中的电势能和动力势能的计算原理进行了类比，认为p指数是一种效能型指标，具有开展综合性学术评价的能力^[15]。Prathap^[14, 16]采用p指数和h指数对100位经济学家、233个国家的科研产出和3个学术机构的学术影响力进行比较研究发现，p指数与h指数之间存在较高的相关性，并且相比于传统指标，p指数能够更有效地平衡被引次数与平均被引率之间的关系，但p指数也存在不能体现论文引文分布特征、不能实现跨领域评价

等不足。故Prathap于2014年引入了被引集中度指标，提出了一个新的综合性评价指标——z指数^[17]。但一些研究表明，z指数并没有起到对p指数优化的作用，z指数对反映数量的被引频次体现较好，但对反映质量的篇均被引(影响因子)和反映一致性的被引集中度体现较弱^[18]。

随着现代科学的蓬勃发展，学科交叉日益深入，跨学科领域开展学术评价的需求不断增加，学界也开展了诸多探索，大致分为两类：一是对被引量排序位置进行处理的跨学科评价方法，细分为以百分位等级为基础的分位法(如Bornmann等^[19]提出的PR6指标、刘雪立等^[20]提出的PR8指标和Leydesdorff等^[21]提出的PR100指标)和以反向名次为基础的位置法(如Pudovkin等^[22]设计的rnf指数和匈牙利科学计量学家Vinkler提出的PRP指数^[23-24])，但此类方法评价结果的稳定性较差，会受期刊分区权重设计和核心期刊目录划定的直接影响。二是对被引量绝对数值进行处理的跨学科评价方法，学者们相继提出了相对被引率(RCR)^[25]、基于主题的标准化被引量期望(NMCR)^[26]、王冠指标(Crown Indicator, CPP/FCSm)^[27]、标准化引文得分期望(MNCS)^[28]等一系列指标；在“引用端”，也有SNCS₍₁₎^[29]、SNCS₍₂₎^[30]、SNIP^[31]、CSNCR^[32]等一系列指标。其计算步骤可大致归纳为四步：①计算某科研实体的篇均被引量；②界定该科研实体所属学科领域；③计算该学科领域论文的平均被引量或平均参考文献量；④基于平均被引，修正篇均被引量^[33]。但需要注意的是，学术影响力应该体现在论文数量和论文质量的有机统一上，上述跨领域评价指标仅仅考虑了对篇均被引量的修正，对发文量的关注始终不足，因此仍有进一步优化空间。

鉴于此，本研究试图以兼顾发文数量和质量为p指数为基础，引入加权影响力调节因子，降低不同学科领域引用平均水平差异所带来的影响，构建一个更为合理的跨领域研究成果评价指标。鉴于该指标能够更加公平合理地对待学科领域之间的差异性，本研

究拟采用公平的英文单词 Fairness 的首字母 F 对指数命名, 称为 F 指数 (Fairness index)。

2 F 指数的构建

2.1 F 指数的计算公式

不同学科领域论文的平均被引次数存在着很大差异, 跨学科的学术影响力评价必须充分考虑这一情况。对于任何一篇文章而言, 若用 k 表示相同学科、相同年份、相同类型论文的平均被引次数, 则可以构造影响力调节因子 $\lambda = \frac{1}{k}$ 。 k 越大, λ 越小, 表示在相同被引数情况下, 该学科领域论文的实际影响力越小。反之亦然。若某篇文章分属两个学科, 则假设两个领域篇均被引值分别是 k_1 和 k_2 , 则这篇文章对应的领域内平均水平 $k = \frac{2}{\frac{1}{k_1} + \frac{1}{k_2}}$ 。分属多个学科的计算方

法同理, 即将相关领域的篇均被引值计算调和平均数, 视为一个全新的小学科领域的平均水平, 以此排除学科差异的影响, 使结果更为客观。对于学者而言, 假设其研究成果共涉及 s 个不重复的学科领域, 每个小学科领域的领域内平均水平分别为 k_1 、 k_2 、 \dots 、 k_s , 该学者在第 i 个小学科领域的论文总被引数为 c_i , 论文总被引数 $C = \sum_{i=1}^s c_i$, 则定义学者的加权影响力调节因子 $\bar{\lambda} = \sum_{i=1}^s \frac{c_i}{C} \cdot \frac{1}{k_i} = \sum_{i=1}^s \omega_i \cdot \lambda_i$ 。该式表示加权影响力调节因子等于小学科领域影响力调节因子的加权平均, 相应的权重 $\omega_i = \frac{c_i}{C}$, 即第 i 个小学科领域总被引数占所有论文总被引数的比例。

若一位学者在考察期内共发表了 n 篇论文, 他的篇均被引为 $\bar{c} = \frac{C}{n}$, 他的加权影响力调节因子为 $\bar{\lambda}$, 则该学者的 F 指数为:

$$F = n^{\frac{1}{3}} \cdot \bar{c}^{\frac{2}{3}} \cdot \bar{\lambda}^{\frac{2}{3}}$$

从计算公式上看, F 指数实现了对学者论文数量、论文影响力和研究领域的统筹兼顾, 即 F 指数会随着发文量 n 、篇均被引 \bar{c} 的增大而增大。同时, 由于 F 指数中包含加权

影响力调节因子 $\bar{\lambda}$, 可调节不同学科领域差异, 从而实现跨领域评价。

2.2 F 指数与相关指标的区别与关联

作为 p 指数的改进, F 指数比 p 指数包含了更多的信息量, 即在遵循 p 指数评价逻辑的基础上, F 指数引入了加权影响力调节因子 $\bar{\lambda}$, 控制了不同学者由于分处不同小学科领域对评价结果带来的影响。p 指数的定义为 $p = \left(\frac{C^2}{n}\right)^{\frac{1}{3}}$, 其中 n 为总发文量, C 为论文总被引次数, 则 $p = \left(\frac{C^2}{n}\right)^{\frac{1}{3}} = n^{\frac{1}{3}} \cdot \bar{c}^{\frac{2}{3}}$, 即得 $F = p \cdot \bar{\lambda}^{\frac{2}{3}}$ 。因此, 相较于 p 指数, F 指数是更为合理的评价指标。

相较于在跨学科评价中使用最广泛的标准引文得分期望 (MNCS) 这一“王冠指标”, F 指数比 MNCS 包含了更多的信息量。

标准化引文得分期望 $MNCS = \frac{\sum_{i=1}^s \frac{c_i}{e_i}}{n}$, c_i 和 n 的定义与前文一致, e_i 表示所属学科论文的平均被引量^[29]。根据 e_i 统计口径的不同, MNCS 也会有所不同。若将相同学科、相同年份、相同类型论文的平均被引次数 k_i 视作所属学科论文的平均被引量 e_i (即令 $e_i = k_i$), 则经简单推导可知 $MNCS = \frac{C}{n} \sum_{i=1}^s \frac{c_i}{C} = \frac{1}{k_i} = \bar{c} \cdot \bar{\lambda}$, 因此 F 指数和 MNCS 的关系为 $F = n^{\frac{1}{3}} \cdot \bar{c}^{\frac{2}{3}} \cdot \bar{\lambda}^{\frac{2}{3}} = n^{\frac{1}{3}} \cdot MNCS^{\frac{2}{3}}$, 说明 F 指数考虑了发文量 n 对学术影响力的影响, 同时发文质量评价指标 MNCS 的次方 ($\frac{2}{3}$) 高于发文数量评价指标 n 的次方 ($\frac{1}{3}$), 也体现出鼓励学者多发高影响力文章的评价导向, 是更为全面的评价指标。

3 F 指数的实证分析

3.1 数据来源及分析过程

本研究基于 Scopus 数据库, 以哲学 (Philosophy in Arts and Humanities)、法学 (Laws in Social Sciences)、统计与概率 (Statistics and

Probability in Mathematic)、机械工程 (Mechanical Engineering in Engineering)、作物学 (Agronomy and Crop Science in Agricultural and Biological Science)、传染病学 (Infectious Diseases in Medicine)、多学科 (Multidisciplinary) 为研究领域, 获取 SciVal 分析平台统计的上述每个领域近五年发文量最多的 500 位学者的相关数据^① (发文量、被引量、FWCI 和 h 指数), 得到 3500 位学者的数据。

由于少数论文获得了绝大部分的引用, 因此引文分布是有偏分布^[34]。每个小学科领域都存在着一批“头部学者”, 其被引总量和 FWCI 均名列前茅。同时, 也存在一批“尾部学者”, 其被引总量和 FWCI 几乎可以忽略不计。若将“头部学者”或“尾部学者”作为研究对象, 各种评价指标会有很强的一致性, 难以判断出 F 指数相较于其他指标在统计性质上的优越性, 故本研究主要关注“中部学者”。

为此, 本研究的分析过程为: ①将每个领域的 500 位学者按被引总量降序排列, 被引总量相同则按 FWCI 降序排列; ②保留每个领域处于上四分位数和下四分位数之间的 250 位学者, 共 1750 位学者; ③将多学科领域的学者数据用于跨学科比较中 F 指数及其他相关指数的统计性质分析; ④将其他六个领域的学者数据用于单一学科比较中 F 指数及其他相关指数的统计性质分析。

3.2 F 指数对跨学科领域学术影响力评价的适用性

由于 h 指数和 p 指数不是进行跨学科比较的评价指标, 因此分析 F 指数在跨学科比较的适用性时, 不必太过关注 F 指数与 h 指数、p 指数的相关性, 而应深入具体学者的发文情况, 评判 F 指数是否起到了调节不同学科领域平均被引次数差异的效果。同时, FWCI 是一个对被引量绝对数值进行处理的跨领域评价指标, 本质上就是以 Scopus 数据库定义的领域内平均水平为 k_i 的标准化引文得分期望 (MNCS)^[35]。因此, 本研究也将学者的 FWCI 排名和 F 指数排名进行比较, 若两者有较高的一致性, 则说明 F 指数具有调节

学科差异开展学术影响力评价的功能。在此基础上, 进而考察 F 指数是否相较于 FWCI 具有良好的区分度, 从而对其跨学科学术影响力评价的适用性作出评判。

3.2.1 对于学科差异的调节作用

多学科领域内的学者主要从事跨学科研究, 且涉及领域大不相同, 是研究跨学科比较的合适样本。表 1 呈现了 F 指数排位前 10 名和后 10 名学者的学术影响力数值。结果表明, h 指数和 p 指数有相对一致的顺序判断, 但均与 F 指数的判断有显著差异, FWCI 虽相较于 h 指数和 p 指数可实现跨学科评价, 但也与 F 指数的判断有显著区别。

相较于 p 指数, F 指数能够更敏锐地探测跨学科研究成果在不同学科影响力的异质性。本研究以 p 指数排名第一的 Launer 和 F 指数排名第一的 Mallapaty 为例, 进行具体分析。两位学者的具体数据如表 2 所示。

通过计算可得, Launer 的篇均被引为 44.5, 而 Mallapaty 的篇均被引仅为 10.2, 与前者存在较大的差距。由此, 似乎 p 指数的排名情况更为合理。但需要注意的是, 在 multidisciplinary 领域, 不同学者涉及的学科领域并不相同, 而这些学科领域的平均被引次数也不相同, 需要深入研究两位学者的被引结构。由于篇幅所限, 表 3 以文章的被引量降序排列, 保留每位作者排名前五的论文被引情况^②。

Launer 排名第一的文章被引量高达 549, 超过 Mallapaty 所有文章的总被引量, 但该文章所归属的学科领域平均被引次数为 15.8, 说明该学科领域普遍有着较高的引用量, 这也稀释了这篇文章的影响力。反观 Mallapaty 排名第一的文章被引量为 98, 但该文章所归属的学科领域平均被引次数仅为 2.2, 说明该学科领域的引用量极低, 该文章被引量为该学科领域平均被引次数的 44.99 倍, 其影响力不言而喻。如果以整体的视角分析两位学者的数据会发现, Launer 涉及的学科领域的

① 检索时间为 2021-08-14, 为了保证统计口径的一致性, 发文量、被引量、FWCI 的统计时间段均为 2016—2020 年, h 指数则是截至 2020 年 12 月 31 日的更新数据。

② 数据来源 SciVal 分析平台, 检索时间为 2021-08-14, 需要全表请与作者联系 wcy1985@tsinghua.edu.cn。

表 1 多学科领域学者的学术影响力测算(部分结果)^①

作者	发文量	被引量	h 指数 (排名)	p 指数 (排名)	FWCI (排名)	F 指数 (排名)
Mallapaty, Smriti	55	561	13 (227)	17.89 (91)	4.58 (3)	10.49 (1)
Ledford, Heidi	69	619	12 (233)	17.71 (97)	3.69 (9)	9.79 (2)
Mahvi, Amir Hossein	23	628	62 (73)	25.79 (43)	5.98 (1)	9.37 (3)
Willyard, Cassandra	24	594	13 (227)	24.50 (53)	4.98 (2)	8.41 (4)
Munafo, Marcus R.	33	731	73 (55)	25.30 (46)	3.89 (6)	7.93 (5)
Smith, George D.	37	1054	209 (1)	31.08 (16)	3.63 (10)	7.87 (6)
Rabinskiy, Lev N.	34	534	28 (187)	20.32 (77)	3.74 (8)	7.81 (7)
Kupferschmidt, Kai	83	907	19 (215)	21.48 (71)	2.32 (33)	7.64 (8)
Cohen, Jon	120	1061	30 (179)	21.09 (72)	1.9 (47)	7.57 (9)
Pham, Viet Thanh	42	842	46 (112)	25.65 (44)	3.11 (14)	7.41 (10)
Zhao, Xingquan Qquan	28	229	36 (154)	12.33 (171)	0.39 (238)	1.62 (241)
Mervis, Jeffrey D.	101	101	13 (227)	4.66 (250)	0.19 (250)	1.54 (242)
Yoshimura, Jin	26	183	20 (210)	10.88 (198)	0.37 (241)	1.53 (243)
Nakayama, Takeo	23	112	39 (146)	8.17 (238)	0.39 (239)	1.52 (244)
Kawashima, Ryuta	30	165	63 (70)	9.68 (218)	0.34 (245)	1.51 (245)
Enshasy, Hesham Ali El	33	117	23 (200)	7.46 (246)	0.31 (246)	1.47 (246)
Sung, Fungchang	25	132	45 (122)	8.87 (228)	0.35 (244)	1.45 (247)
Cressey, Daniel	54	151	16 (220)	7.50 (244)	0.23 (249)	1.42 (248)
Wang, Anxing Xin	27	185	25 (196)	10.82 (200)	0.31 (247)	1.37 (249)
Lee, Jong-min	24	135	46 (112)	9.12 (225)	0.31 (248)	1.32 (250)

表 2 Launer 和 Mallapaty 在多学科领域的相关指标表现

作者	发文量	被引量	h 指数 (排名)	p 指数 (排名)	FWCI (排名)	F 指数 (排名)
Launer	24	1069	151 (7)	36.24 (1)	3.03 (17)	6.04 (21)
Mallapaty	55	561	13 (227)	17.89 (91)	4.58 (3)	10.49 (1)

^①本表仅列出前 10 名和后 10 名学者的指标数值,需要全表请与作者联系 wcy1985@tsinghua.edu.cn.

表 3 Launer 和 Mallapaty 在多学科领域的部分论文被引情况(2016—2020)

作者	文章名	被引量	FWCI	k
Launer	Genome-wide association study identifies 74 loci associated with educational attainment	549	23.83	15.8
	Co-regulatory networks of human serum proteins link genetics to disease	100	6.2	16.1
	The genetic architecture of the human cerebral cortex	98	17.18	5.2
	1000 Genomes-based meta-analysis identifies 10 novel loci for kidney function	51	2.71	17
	Genome-Wide Association Study for Incident Myocardial Infarction and Coronary Heart Disease in Prospective Cohort Studies: The CHARGE Consortium	50	2.21	15.8
Mallapaty	The coronavirus pandemic in five powerful charts	98	44.99	2.2
	How sewage could reveal true scale of coronavirus outbreak	90	41.31	2.2
	The coronavirus is most deadly if you are older and male-new data reveal the risks	48	22.03	2.2
	What the cruise-ship outbreaks reveal about COVID-19	40	18.36	2.2
	How China could be carbon neutral by mid-century	34	15.61	2.2

平均被引次数偏高, 24 篇文章中有 14 篇文章的领域平均被引次数 > 10, 有 10 篇文章的领域平均被引次数 > 15, 而 Mallapaty 涉及的学科领域的平均被引次数偏低, 55 篇文章中有 43 篇文章的领域平均被引次数 < 3。因此, 若仅仅以学者的发文量和总被引量的数据对学者进行评价, 而不考虑其所在学科领域的相关信息, 就会埋没一批在领域平均被引次数小的学科领域中独占鳌头的学者。F 指数能够较好地解决这一问题, 对此类学者予以关注。

相较于 FWCI, F 指数与其相关系数高达

0.943, 存在显著正相关, 说明 F 指数与 FWCI 一样, 均具有跨领域评价的功能。例如, 如表 4 所示, 以 F 指数排名第一的 Mallapaty 和 FWCI 排名第一的 Mahvi, 均属于高学术影响力的第一梯队, 但 Mallapaty 的发文量 (55) 远超 Mahvi 的发文量 (23), 这意味着在学术影响力相当的基础上, 前者有着发文数量的绝对优势。因此就学术影响力而言, Mallapaty 应大于 Mahvi, 而 F 指数正是考虑了学术产出数量的因素, 所以能够得到更为科学合理的评价结果。

表 4 Mahvi 和 Mallapaty 在多学科领域的相关数据汇总

作者	发文量	被引总量	h 指数 (排名)	p 指数 (排名)	FWCI (排名)	F 指数 (排名)
Mallapaty	55	561	13 (227)	17.89 (91)	4.58 (3)	10.49 (1)
Mahvi	23	628	62 (73)	25.79 (43)	5.98 (1)	9.37 (3)

3.2.2 评价结果的区分度

一方面, F 指数可对不同学科学术影响力的异质性进行调节, 从而实现跨学科研究成果的学术影响力评价; 另一方面, F 指数必须在评价结果上具有更加精准的区分度, 才能优于现有指标。本研究认为, 合理的区分度要求每个分数段内的人数尽可能接近正态分布, 且同分 (无法判断) 的人数尽可能少, 以免降低评价效率。

从正态性检验来看, F 指数是最为接近正态分布的评价指标。根据 Kolmogorov-Smirnov 检验, h 指数的 D 统计量为 0.12777, p 值为 5.699×10^{-4} ; p 指数的 D 统计量为 0.12955, p 值为 4.533×10^{-4} ; FWCI 的 D 统计量为 0.16167, p 值为 4.218×10^{-6} ; F 指数的 D 统计量为 0.11123, p 值为 4.118×10^{-3} 。虽然它们都拒绝了数据分布符合正态分布的零假设, 但从 D 统计量和 p 值可以看出, F 指数最

接近正态分布, FWCI 最不接近正态分布。

从同分情况来看, F 指数出现同分的情况要显著低于 h 指数和 FWCI。表 5 显示 250 名学者中, h 指数有 43 名学者不同分, p 指数有 224 名学者不同分, FWCI 有 89 名学者不同分, F 指数有 148 名学者不同分。其中相较于同样具有跨领域评价的 FWCI, F 指数同分 2 名及以上学者的人数, 显著要低。

表 5 多学科领域的同分学者数量分布表

领域	同分人数	h 指数	p 指数	FWCI	F 指数
多学科	0	43	224	89	148
	1	52	26	60	54
	2	42	0	54	30
	3—5	96	0	47	18
	6 以上	17	0	0	0

表 6 F 指数及参照指标相关系数汇总表

	发文量	篇均被引	h 指数	p 指数	FWCI
哲学	-0.023	0.62***	0.16*	0.65***	0.97***
法学	-0.039	0.72***	0.23***	0.77***	0.97***
统计与概率	-0.21***	0.78***	0.11*	0.82***	0.95***
机械工程	-0.038	0.54***	0.36***	0.59***	0.94***
作物学	-0.082	0.56***	0.18**	0.61***	0.96***
传染病学	-0.063	0.71***	0.12*	0.78***	0.93***

更少的同分情况。

六个学科领域中 250 位学者的 h 指数、p 指数、F 指数和 FWCI 经过 Kolmogorov-Smirnov 检验的结果显示^①(见表 7)。在各个学科领域, F 指数的 D 统计量均为最小值, 这表明 F 指数相较于其他指数更接近正态分布; 就显著性水平而言, 除了哲学领域之外, F 指数在其他学科领域都近似于正态分布。此外, 在大多数学科领域中, p 指数、FWCI 和 h 指数的 D 统计量逐渐增大, 这代表着 p 指数是除 F 指数以外, 更接近正态分布的评价指标, 但也仅在机械工程领域近似于正态分布。

六个学科领域中 250 位学者的 h 指数、p 指数、FWCI 和 F 指数的同分情况显示(见表 8), F 指数和 p 指数的同分人数小于 h 指数和 FWCI, 哲学领域 F 指数的同分人数小于 p 指数, 但在其他领域均略低于 p 指数。需要强调的是, 由于 p 指数和 F 指数都涉及到开方的运算, 因此在保留更多位有效数字的情况下,

综上, 不同小学科领域的发文量差异会对学者的学术影响力评价结果产生显著影响, F 指数是更为合理和全面的跨学科评价指标。

3.3 F 指数在单一学科领域学术影响力评价的适用性

在单一学科学术影响力的评价中, h 指数和 p 指数虽具有一定的局限性, 但也被广泛应用。F 指数与 h 指数、p 指数均具有显著的相关性(见表 6), 特别是与 p 指数在各学科领域均呈现中度以上的正向显著相关。这表明, F 指数并非是对现有评价指标的颠覆, 而是进一步的改进和完善, 可用于单一学科学术影响力的评价。

进一步, 本研究将考察 F 指数在区分度上的表现, 是否更接近正态分布、以及是否出现

F 指数和 p 指数几乎都不会出现同分的情况, 从此意义上讲, F 指数的区分度能够得到有效保证。

4 结语

本研究提出 F 指数作为一种改进的学术影响力测度方法, 并分别以 2016—2020 年间在多学科、哲学、法学、统计与概率、机械工程、作物学、传染病学等七个领域各 250 位学者为研究对象, 进行了实证研究。结果发现, F 指数虽与 h 指数、p 指数、FWCI 显著相关, 但能够实现跨学科领域的学术影响力比较, 且具有更高的区分度, 其分布更接近正态分布, 是适用性更广、科学性更强的评价指标。F 指数也可用于学者、科研团队及学术组织的学术影响力评价。

^① 显著性一列中, *** p<0.001; ** p<0.01; * p<0.05; + p<0.1, 未标明显著性水平的指数则接受 Kolmogorov-Smirnov 检验的零假设, 认为该指数的分布为正态分布。

表 7 正态分布检验的 D 统计量数值及 p 值

学科领域	h-指数	p 指数	FWCI	F 指数	D 统计量比较	显著性水平
哲学	$D_1 = 0.16531$ $\rho_1 = 2.33 \times 10^{-6}$	$D_2 = 0.092848$ $\rho_2 = 0.02686$	$D_3 = 0.12829$ $\rho_3 = 0.0005338$	$D_4 = 0.086333$ $\rho_4 = 0.04814$	$D_4 < D_2 < D_3 < D_1$	h 指数*** FWCI*** p 指数* F 指数*
法学	$D_1 = 0.15706$ $\rho_1 = 8.794 \times 10^{-6}$	$D_2 = 0.078465$ $\rho_2 = 0.09206$	$D_3 = 0.062696$ $\rho_3 = 0.2795$	$D_4 = 0.047181$ $\rho_4 = 0.6339$	$D_4 < D_3 < D_2 < D_1$	h 指数*** p 指数+
统计与概率	$D_1 = 0.13112$ $\rho_1 = 0.0003698$	$D_2 = 0.07978$ $\rho_2 = 0.08297$	$D_3 = 0.088001$ $\rho_3 = 0.04163$	$D_4 = 0.056153$ $\rho_4 = 0.4097$	$D_4 < D_2 < D_3 < D_1$	h 指数*** FWCI* p 指数+
机械工程	$D_1 = 0.12749$ $\rho_1 = 0.0005911$	$D_2 = 0.06571$ $\rho_2 = 0.2306$	$D_3 = 0.06751$ $\rho_3 = 0.2046$	$D_4 = 0.059655$ $\rho_4 = 0.3359$	$D_4 < D_2 < D_3 < D_1$	h 指数***
作物学	$D_1 = 0.10909$ $\rho_1 = 0.005208$	$D_2 = 0.084898$ $\rho_2 = 0.05443$	$D_3 = 0.089183$ $\rho_3 = 0.03749$	$D_4 = 0.058804$ $\rho_4 = 0.353$	$D_4 < D_2 < D_3 < D_1$	h 指数** FWCI* p 指数+
传染病学	$D_1 = 0.073093$ $\rho_1 = 0.1383$	$D_2 = 0.090721$ $\rho_2 = 0.03265$	$D_3 = 0.12097$ $\rho_3 = 0.001329$	$D_4 = 0.067998$ $\rho_4 = 0.198$	$D_4 < D_1 < D_2 < D_3$	FWCI** p 指数*

表 8 分领域的同分学者数量分布表

同分人数	领域	h 指数	p 指数	FWCI	F 指数	领域	h 指数	p 指数	FWCI	F 指数
0	哲学	3	89	77	126	法学	13	157	103	144
1		10	56	66	72		12	74	82	78
2		6	54	57	39		6	15	39	24
3-5		15	44	43	13		28	4	26	4
6 以上		216	7	7	0		191	0	0	0
0	统计与概率	13	167	47	94	机械工程	16	201	60	126
1		10	68	76	86		26	46	72	74
2		27	15	57	48		15	3	60	33
3-5		84	0	70	22		56	0	58	17
6 以上		116	0	0	0		137	0	0	0
0	作物学	13	202	54	100	传染病学	18	203	64	113
1		10	38	72	86		36	32	52	76
2		24	6	45	48		30	15	63	39
3-5		69	4	71	16		82	0	71	22
6 以上		134	0	8	0		84	0	0	0

但 F 指数也存在着继续完善的空间。通过理论推导和具体比较几位学者 p 指数与 F 指数、FWCI 与 F 指数的评价结果，仅是初步证明 F 指数可实现跨领域评价，如何能够提升 F 指数在跨领域比较的评价效能仍需进一步思考，也可考虑引入更多用于跨领域评价的指

标，如影响因子百分位、标准化特征因子^[36-38]等。此外，囿于数据的可得性，研究数据的被引总量、FWCI 均包含学者自引，如能除去自引，其评价结果会更为合理，但自引同时也影响 h 指数、p 指数和 FWCI 的数值表现，故整体上不会影响本研究主要结论。

参考文献

[1] 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.

[2] Rousseau R, 李志忠. h 指数研究在 2007—2008 年的新进展[J]. 科学观察, 2008, 3(3): 19-25.

[3] 聂超, 袁浩川, 高慧颖. 一种改进 h 指数的方法[J]. 情报资料工作, 2009(6): 84-87.

[4] Hirsch J E, Buela-Casal G. The meaning of the h-index[J]. International Journal of Clinical and Health Psychology, 2014(14): 2, 161-164.

[5] 宋振世, 周健. 类 h 指数研究综述及其实证对比分析[J]. 情报科学, 2012, 30(11): 1652-1657.

[6] 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.

- [7] Egghe L. Theory and practice of the g-index[J]. *Scientometrics*, 2006, 69(1): 131-152.
- [8] 金碧辉. 科学家为自己设计了一项评价指标: h 指数[J]. *科学观察*, 2006, 1(1): 8-9.
- [9] 金碧辉, Rousseau R. R 指数、AR 指数: h 指数功能扩展的补充指标[J]. *科学观察*, 2007, 2(3): 1-8.
- [10] Alonso S, Cabrerizo F J, Herrera Viedma E, et al. Hg-index: A new index to characterize the scientific output of researchers based on the h- and g-indices[J]. *Scientometrics*, 2010, 82(2): 391-400.
- [11] 聂超, 魏泽峰. 基于学术影响力差异 h 指数改进的实证研究[J]. *情报杂志*, 2010, 29(5): 89-91.
- [12] 王凌峰. 一个新的 h-type 指标——A+ 指数[J]. *情报杂志*, 2013, 32(1): 55-58.
- [13] 张学梅. h_1 指数——一种对 h 指数进行迭代计算的学术影响力评价方法[J]. *图书情报工作*, 2013, 57(11): 96-99, 119.
- [14] Prathap G. Is there a place for a mock h-index? [J]. *Scientometrics*, 2010, 84(1): 153-165.
- [15] Prathap G, Nishy P. E = π^2 -The energy of ideas approach to bibliometric research assessment[J]. *Annals of Library and Information Studies*, 2010, 57(3): 282-286.
- [16] Prathap G. The 100 most prolific economists using the p-index[J]. *Scientometrics*, 2010, 84(1): 167-172.
- [17] Prathap G. The zynergy-index and the formula for the h-index[J]. *Journal of the Association for Information Science and Technology*, 2014, 65(2): 426-427.
- [18] 俞立平, 王作功. z 指数评价学术期刊的适用性及其改进研究[J]. *情报学报*, 2018, 37(11): 1132-1139.
- [19] Bornmann L, Mutz R. Further steps towards an ideal method of measuring citation performance: The avoidance of citation (ratio) averages in field-normalization[J]. *Journal of Informetrics*, 2011, 5(1): 228-230.
- [20] 刘雪立, 魏雅慧, 盛丽娜, 等. 期刊 PR8 指数: 一个新的跨学科期刊评价指标及其实证研究[J]. *图书情报工作*, 2017, 61(11): 116-123.
- [21] Leydesdorff L, Bornmann L, Mutz R, et al. Turning the tables on citation analysis one more time: Principles for comparing sets of documents[J]. *Journal of the American Society for Information Science and Technology*, 2011, 62(7): 1370-1381.
- [22] Pudovkin A I, Garfield E. Rank-normalized impact factor: A way to compare journal performance across subject categories[J]. *Proceedings of the American Society for Information Science and Technology*, 2004, 41(1): 507-515.
- [23] Vinkler P. Comparative rank assessment of journal articles[J]. *Journal of Informetrics*, 2013, 7(3): 712-717.
- [24] 李华, 张学梅, 刘文芝. PRP 指标: 一种跨学科期刊评价方法及与期刊 H 指数之比较[J]. *情报杂志*, 2015, 34(8): 60-64, 109.
- [25] Schubert A, Glänzel W, Braun T. Relative citation rate: A new indicator for measuring the impact of publications[C]// *Proceedings of the First National Conference with International Participation on Scientometrics and Linguistics of Scientific Text*, 1983: 80-81.
- [26] Braun T, Glänzel W. United Germany: The new scientific superpower? [J]. *Scientometrics*, 1990, 19(5-6): 513-521.
- [27] van Raan A F J. Measurement of central aspects of scientific research: Performance, interdisciplinarity, structure [J]. *Measurement Interdisciplinary Research and Perspectives*, 2005, 3(1): 1-19.
- [28] Waltman L, van Eck N J, van Leeuwen T N, et al. Towards a new crown indicator: Some theoretical considerations [J]. *Journal of Informetrics*, 2011, 5(1): 37-47.
- [29] Zitt M, Small H. Modifying the journal impact factor by fractional citation weighting: The audience factor[J]. *Journal of the American Society for Information Science and Technology*, 2008, 59(11): 1856-1860.
- [30] Waltman L, van Eck N J. A systematic empirical comparison of different approaches for normalizing citation impact indicators[J]. *Journal of Informetrics*, 2013, 7(4): 833-849.
- [31] Moed H F. Measuring contextual citation impact of scientific journals[J]. *Journal of Informetrics*, 2010, 4(3): 265-277.
- [32] Bornmann L, Haunschild R. Citation score normalized by cited references (CSNCR): The introduction of a new citation impact indicator[J]. *Journal of Informetrics*, 2016, 10(3): 875-887.
- [33] 张学梅. 科学计量学视域下跨学科学术评价方法研究进展[J]. *情报理论与实践*, 2020, 43(7): 171-178.
- [34] Seglen P O. The skewness of science[J]. *Journal of the American Society for Information Science*, 1992, 43(9): 628-638.
- [35] 吴伟, 姜天悦, 余敏杰. 我国高水平大学基础研究与世界一流水平的群体性差距——基于学科规范化的引文影响力分析[J]. *现代教育管理*, 2017(4): 18-23.
- [36] 顾欢. 影响因子百分位与 SNIP 的跨学科评价效力实证分析[J]. *情报杂志*, 2017(7): 134-137.
- [37] 牛晓锋. 不同影响因子百分位对期刊评价效力的比较研究[J]. *出版广角*, 2018(12): 52-54.
- [38] 方红玲. 特征因子百分位用于期刊跨学科评价的可行性研究[J]. *中国科技期刊研究*, 2020, 31(7): 845-849.

(收稿日期: 2021-09-15)