

不同学科期刊学术影响力比较的方法与实证研究

■ 陈福佑 杨立英 丁洁兰

[摘要] 对不同学科期刊学术影响力比较中的常见问题进行详细描述,分析问题产生的根源,在此基础上系统梳理现有的解决方案,最后从实证角度以学科归一化的期刊评价指标相对期刊影响因子(relative impact factor, RIF) 和平均百分比等级指标(mean percentile rank, MPR) 为例,引入“公平性测试”(fairness test) 方法来检验这两个学科归一化评价指标在不同学科比较中的适用性。

[关键词] 期刊学术影响力 学科归一化 RIF MPR 公平性测试

[分类号] G350

DOI: 10. 7536/j. issn. 0252 - 3116. 2013. 23. 014

在利用文献计量方法评价科学活动的过程中,经常会遇到对不同学科进行比较的需求,例如评价不同学科的发展水平,对比不同学科科研机构的绩效。学科之间比较的难点在于不同学科由于研究规模差异、研究性质不同,其相应的文献计量评价指标得分往往差距较大,因此,不可以直接进行对比。

期刊学术影响力评价是期刊评价乃至科研评价的基础内容之一。在评价实践中,常常使用文献计量指标作为辅助依据。与其他评价对象类似,不同学科之间比较的问题也同样存在于期刊评价之中。典型的例子是一些常用的期刊评价指标,如影响因子,有着显著的学科特征,不同学科期刊的影响因子得分差异很大。若直接使用影响因子进行不同学科期刊的影响力分析,则往往会发现某学科期刊的影响因子指标普遍高于或低于另一个学科,这说明影响因子这类学科间差异较大的指标不适用于进行不同学科之间的比较。

在文献计量领域,研究人员已经提出不同学科之间期刊比较的若干实践解决方案。针对中文学术期刊的评价,我国研究人员提出了若干解决方案,主要集中在对不同学科的期刊影响因子调整方面,对于其他期刊评价指标的研究相对较少^[1-7]。因此,无论是国际还是国内,相关的理论研究仍有待加强。此外,各种评价指标归一化处理后的效率对比也较为缺乏。本文试从理论上梳理不同学科期刊学术影响力比较问题的具体表现、产生根源与解决方案,同时,引入“公平性测试”方法对不同学科的期刊归一化指标进行实证分析,

以丰富现有的期刊学术影响力指标比较研究。

1 不同学科期刊学术影响力比较概述

1.1 不同学科期刊学术影响力比较中的常见问题

在进行期刊学术影响力学科比较时,使用非学科归一化指标直接对不同学科的期刊进行评价,会出现“学科偏见”(discipline-dependent bias)的问题^[8],例如对不同学科期刊的影响因子直接对比,会使得一个学科的得分普遍位于另外一个学科之下,影响评价的公平性。

图 1 以汤森路透 JCR(*Journal of Citation Report*) 数据库发布的期刊影响因子数据为例,选取 2011 年数学(Mathematics) 和肿瘤学(Oncology) 两个学科为样本来说明用影响因子进行学科间比较存在的问题。从图 1 可以看出:整体而言,数学学科的绝大多数期刊影响因子低于肿瘤学。肿瘤学领域影响力最高的期刊 *CA: A Cancer Journal for Clinicians*,其影响因子超过 100;在数学领域,影响力最高的期刊 *Journal of the American Mathematical Society* 只有 3.841,如果把这本期刊直接与肿瘤学领域期刊比较,在肿瘤学期刊中排名第 52 位。因此,通过直接比较两个学科期刊的影响因子来评价两个学科期刊的学术影响力水平,对于数学学科显然不公平。

1.2 不同学科期刊学术影响力比较中常见问题产生的根源

究其根源,是因为在进行期刊学术影响力学科间

[作者简介] 陈福佑,中国科学院大学、中国科学院国家学科图书馆硕士研究生,E-mail: chenfuyong@mail.las.ac.cn; 杨立英,中国科学院国家科学图书馆副研究员,博士; 丁洁兰,中国科学院国家科学图书馆研究实习员,硕士。

收稿日期: 2013 - 10 - 21 修回日期: 2013 - 11 - 20 本文起止页码: 85 - 89 94 本文责任编辑: 王善军

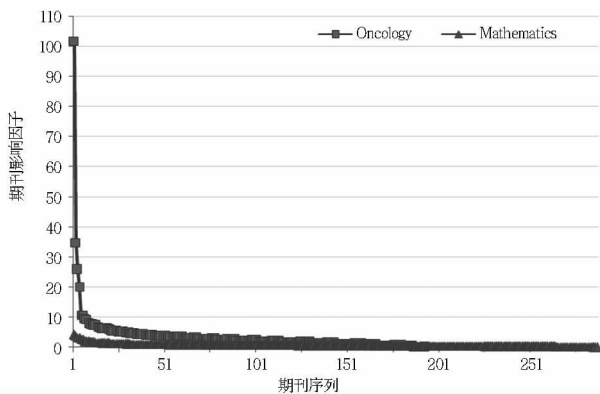


图1 肿瘤学(Oncology)与数学(Mathematics)所有期刊2011年期刊影响因子分布

注:学科分类与学科下的期刊列表与影响因子来源于汤森路透的JCR 2011。

比较时,不能全面认识各学科研究特点、研究性质差异对期刊评价指标的影响,没有合理对评价指标进行学科归一化处理,造成评价指标的不当使用。

各学科研究特点不同,导致引用行为存在差异,例如有的学科研究需要大量借鉴前人的工作,以支撑后者的研究成果,所以发表论文的参考文献数量多,表现出引用活跃度高、引用周期短的特征;有的学科研究工作则相对独立,对前人成果的介绍相对较少,所以参考文献数量相对较少,表现出引用活跃度较低、引用周期较长的特征。

E. Garfield很早就注意到学科引用行为差异问题:不同学科论文的参考文献数量显著不同。例如数学领域的论文参考文献一般在10篇以内,而大多数生物科学的参考文献超过30篇,这就是生物学论文的篇均被引频次远大于数学论文的主要原因^[9]。E. Garfield还明确指出:在文献计量研究中,引用数据相对学科和研究领域来说非常敏感,使用时需要根据学科引用差异做出相应的调整^[10]。

2 不同学科期刊学术影响力比较的方法与指标

2.1 不同学科期刊学术影响力比较的方法

在文献计量研究中,解决不同学科期刊影响力比较的基本思路是处理学科间计量指标的量纲差异。目前主要的解决方案是使用学科归一化的方法对评价指标进行处理,即比较指标与基准线(baseline),将指标的绝对分值转化为相对于基准线的分值^[11-15];除学科归一化外,还有许多其他方法,如SNIP采用期刊主题领域引文潜力对论文原始影响力归一化,MPR则使

用percentile rank依据论文等级赋予不同的权值实现归一化等。M. Zitt和H. Small将归一化方法分为引证归一化(citing normalization)和被引归一化(cited normalization)两大类^[16],而H. F. Moed将归一化方法分为引用源归一化方法(source normalization)和引用目标归一化方法(target normalization)^[17]。本文依据H. F. Moed的分类体系来梳理跨不同学科期刊学术影响力比较方法。

2.1.1 引用源归一化方法 引用源归一化方法从论文引用源(参考文献)的角度出发,从不同学科对比问题产生的根源出发,考虑不同学科平均参考文献数量差异,并将此作为归一化处理的原则。例如,获得相同被引频次的论文,在平均参考文献数量较多的学科,其学术影响力比在平均参考文献数量较少的学科要小。因此,引用源归一化方法根据不同学科参考文献数量的相对水平对期刊论文的被引频次归一化处理,以此消除学科差异^[11,16]。代表指标有SNIP(source normalization impact per paper)、audience factor、fractional citation counting、mean source-normalized citation score^[12]。

2.1.2 引用目标归一化方法 引用目标归一化方法从论文被引频次的角度出发,把不同学科平均被引频次的量纲差异作为归一化处理的原则。例如,获得相同被引频次的论文,在平均被引频次较高的学科,其学术影响力比在平均被引频次少的学科要小。该方法通常使用整个学科的评价指标值(均值或中位数)作为基准线,用基础评价指标除以学科基准线来得到归一化后的得分,如学科相对期刊影响因子(relative impact factor, RIF)、CPP/FCSm(the old “crown indicator”)、citation z-score等^[11,17]。

相对于通常的引用目标归一化方法,平均百分比等级指标(mean percentile rank, MPR)是一种特殊形式的归一化指标,将每一个学科分成若干等级,认为同等级内论文的学术影响力相同,即便是不同学科的论文学术影响力也相同,对它们赋予相同的权值^[18-19],以此达到归一化的目标。

2.2 常用的不同学科期刊学术影响力比较指标

2.2.1 SNIP指标 SNIP是引用源归一化方法的代表性指标之一,它从被评价期刊的施引文献出发,基于施引文献的篇均参考文献数量对被评价期刊进行归一化处理。具体的做法是:将施引期刊视为被评价期刊涉及的主题领域,以主题领域(施引期刊)的篇均参考文献数作为该主题领域的引文潜力(database citation

potential ,DCP) ,用来描述被评价期刊收到被引频次的期望值。在引文数据库,将期刊的 DCP 与数据库中 DCP 的中位数比较,求出相对主题领域的引文潜力 (relative database citation potential ,RDCP) ,RDCP 反映期刊的期刊评价指标与学术影响力的差别,这些差别受其他因素的影响,其中包括学科差异。

SNIP 使用 RDCP 来归一化期刊原始影响力 (Raw impact per paper ,RIP ,即3年篇均被引频次) ,以此来消除学科等因素带来的差异。计算公式如下:

$$SNIP_j = \frac{RIP_j}{R_j^{db}/M^{db}}$$

其中 RIP_j 表示期刊原始影响力, R_j^{db} 是期刊主题领域引文潜力, M^{db} 为数据库中期刊主题领域引文潜力的中位数^[16]。

2.2.2 学科相对期刊影响因子 RIF 是引用目标归一方法的代表性指标,用于测度期刊影响因子 (IF) 在所属学科的相对水平。其基本思路是将学科全部期刊的平均影响因子作为基准线,计算该期刊与基准线的比值,这样就将指标的绝对分值转化为相对于基准线的分值^[2,16]。

RIF 指标的计算公式如下:

$$RIF_x = \frac{IF_x}{\sum IF/n}$$

IF_x 表示某期刊的影响因子, $\sum IF/n$ 表示某期刊所属学科(全部期刊)的平均影响因子, RIF_x 就是某本期刊的相对影响力水平。

2.2.3 平均百分比等级指标 MPR 是引用目标归一方法的特殊方法,它从期刊被引用的情况出发,通过将学科内论文根据被引频次降序排列分成若干个等级,对位于相同等级的论文,即使是不同学科的论文,赋予相同的权值。然后通过计算期刊权值的平均值来评价期刊的学术影响力。MPR 不仅能够消除学科给期刊学术影响力带来的影响,而且能够解决期刊论文被引频次偏态分布造成均值评价指标不能正确评价期刊学术影响力的问题^[18-19]。

MPR 的等级有多种选择,常用的有6等级和100等级。分为6等级的,即 MPR(6),其评价过程如下:首先,将学科内的论文按照被引频次降序排列,根据论文数量 top 1%、top 5%、top 10%、top 25%、top 50% 和 botto50%,划分为6个等级,并定义第1等级到第6等级的论文权值 (X_i) 分别为 6、5、4、3、2、1,然后计算期刊所有论文的平均权值。100等级的,即 MPR(100),则是将期刊的所有论文分为100等份,赋予的权值从100到1^[19]。

3 不同学科期刊影响力比较指标的“公平性”检验

F. Radicchi 和 L. Ledeydorff 提出“公平性测试”方法,从概率的角度为学术期刊评价指标的学科间比较的公平性提供可计量的评价依据。两人利用“公平性测试”方法,实证分析在学科的 top 10% 论文中,归一化指标 fractional citation 的公平性要远远高于期刊影响因子^[8,11]。

在不同学科期刊学术影响力比较的两种方法中,引用目标归一方法更为常见。在引用目标归一方法中,RIF 应用历史较长,应用范围较广泛;MPR 提出相对较晚,是近年来受到更多关注和研究的新兴指标。本文也将利用“公平性测试”方法测试 RIF、MPR、IF 的不同学科比较的公平性。因为学科内不同等级的期刊学术影响力的分布有较大的差异,不同等级的特征也就不相同,所以 L. Ledeydorff 提出测试 top 10% 论文的“公平性”,即便此部分论文的评价指标公平性很高,也不能代表该评价指标就能直接实现不同学科间的比较。本文将在 top 10% 论文的基础上,按照 top 10% - top 20%、top 20% - top 40%、top 40% - top 60%、top 60% - top 80%、top 80% - top 100% 的区间测试不同评价指标的公平性,以此较全面地比较常用的评价指标 RIF、MPR 在不同学科期刊比较的适用性。

3.1 “公平性测试”方法概述

F. Radicchi 和 C. Castellano 提出“公平性”评价指标,认为指标应该不受学科或研究领域影响,即指标得分高低的出现概率与学科无关^[20]。因此,无论在哪个学科,其指标分值高低的概率相同。满足“公平性”原则的评价指标分值在不同学科间没有显著性差异。

L. Leydesdorff 和 F. Radicchi 根据“公平性”评价指标的原理,提出期刊评价指标的“公平性测试”方法,用来查看其受学科影响的程度^[21]。具体方法如下:假设一个数据集包含 G 个学科的 N 篇期刊,其中 N_g 表示属于第 g 个学科的期刊数量。期刊评价指标会给每本期刊一个评价价值,把所有期刊按照评价价值降序排列,从上述排列中抽取 top z% 的期刊, top z% 期刊数量为 $n^z = [zN/100]$ 。 $m_g^{(z)}$ 表示 g 学科在 top z% 中的期刊数量, top z% 期刊在特定学科的出现满足超几何分布,学科期刊在 top z% 期刊中出现的概率如下:

$$P(m_g^{(z)} / n^{(z)} | N, n_g) = \binom{N - N_g}{n^{(z)} - m_g^{(z)}} / \binom{N}{n^{(z)}}$$

其中 $\binom{x}{y} = x! / [y! (x - y)!]$ 。如果期刊评价指标

满足公平性原则,则每个学科 top z% 出现概率都在理想概率区间内^[21]。

3.2 数据集构建

在学科体系中,数学和生命科学是两个学科性质和研究特点迥然相异的学科,从文献计量研究的视角看,由于学科研究性质的差别,两个学科的引用规律差别也很大(见图1),因此本研究选择数学、生命科学作为比较指标公平性测试的样本学科。

本文采纳汤森路透公司 JCR 数据库的学科分类体系,选择与数学、生命科学相关的 Mathematics、Logic、Oncology、Developmental Biology 4 个学科所有期刊 2009 - 2010 年文献类型为 Articles、Proceedings papers、Reviews、Letters 的论文作为样本数据。为保持评价指标计算的一致性,从 WOS(Web of Science) 的 citation report 下载每篇论文在 2011 年的被引频次,然后对数据进行规范,且排除论文数量 5 以下的期刊,最后根据论文的被引频次计算出每本期刊的 IF、RIF、MPR(6) 和 MPR(100) 的数值,以形成公平性测试的数据集。

3.3 “公平性测试”结果分析

将学科的所有论文按照各指标计算后的结果的 20% 的间距,分为 6 等分(因为文献计量数据几乎都是呈长尾分布,0 - 20% 属于头部地区,期刊之间的得分差异相对较大,所以把 0 - 20% 进一步分成 2 个区间),使用“公平性测试”方法计算出 IF、RIF、MPR(6)、MPR(100) 4 个评价指标分别在 Mathematics、Logic、Oncology、Developmental Biology 4 个学科的 [0, 10%]、[10, 20%]、[20%, 40%]、[40%, 60%]、[60%, 80%]、[80%, 100%] 这 6 个等级区间内的超几何分布概率。

表 1 是 IF、RIF、MPR(6)、MPR(100) 在 Mathematics、Logic、Oncology、Developmental Biology 中 6 个等级的实际出现概率(超几何分布概率),其中灰色部分表示实际出现概率在理想概率区间内,即是满足公平性测试的测试结果。图 2(组图)是表 1 的图形化形式,灰色的矩形框代表学科的理想概率区间,各个点则是不同评价指标在不同学科各等级区间的期刊的实际出现概率,代表指标实际出现概率的点如果落在灰色块以内,表示该指标符合“公平性”要求,否则不符合。

表 1 IF、RIF、MPR(6) 和 MPR(100) 在 4 个学科的 6 个等级区间的超几何分布概率(置信度为 90%)

等级	学科	IF	RIF	MPR(6)	MPR(100)
[0 - 10%]	Mathematics	0.000	0.089	0.049	0.005
	Logic	0.147	0.288	0.147	0.303
	Oncology	0.000	0.012	0.115	0.077
	Developmental Biology	0.000	0.023	0.009	0.003
[10% - 20%]	Mathematics	0.000	0.032	0.002	0.019
	Logic	0.147	0.000	0.068	0.288
	Oncology	0.000	0.000	0.006	0.104
	Developmental Biology	0.000	0.030	0.030	0.005
[20% - 40%]	Mathematic	0.000	0.000	0.000	0.000
	Logic	0.080	0.032	0.011	0.032
	Oncology	0.000	0.000	0.000	0.000
	Developmental Biology	0.009	0.102	0.015	0.015
[40% - 60%]	Mathematic	0.000	0.075	0.006	0.001
	Logic	0.219	0.078	0.219	0.219
	Oncology	0.001	0.081	0.033	0.002
	Developmental Biology	0.000	0.095	0.033	0.099
[60% - 80%]	Mathematic	0.000	0.019	0.079	0.000
	Logic	0.032	0.234	0.080	0.174
	Oncology	0.000	0.034	0.049	0.000
	Developmental Biology	0.000	0.092	0.019	0.036
[80% - 100%]	Mathematic	0.000	0.000	0.000	0.000
	Logic	0.081	0.171	0.171	0.171
	Oncology	0.000	0.000	0.000	0.000
	Developmental Biology	0.000	0.033	0.064	0.146

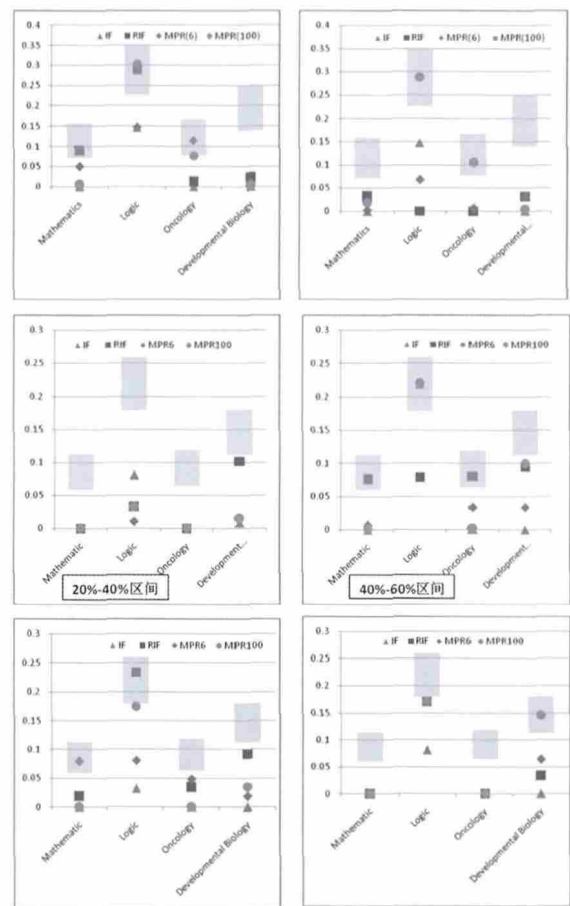


图 2 IF、RIF、MPR(6) 和 MPR(100) 在 4 个学科的 6 个等级区间的超几何分布概率与理想概率区间的对比(置信度为 90%)

从表1和图2可以看出,在“公平性测试”中,IF因为没有经过学科归一化,在每个测试学科、各个区间都没有达到“公平性”的要求。RIF、MPR(6)、MPR(100)满足“公平性”的学科区间也不多。RIF在[0-10%]区间表现得最好,但是也只有两个学科的概率达到公平性要求,而在[10%-20%]、[20%-40%]和[80%-100%]区间的概率都没有达到公平性要求。MPR按照两种等级计算的评价值最终的效果也不同,MPR(100)相对MPR(6)要好一点,但是并不明显。MPR(100)在[0-10%]、[10%-20%]均有两个学科的概率达到了公平性要求,并且[40%-60%]、[60%-80%]均有一个学科达到了公平性要求,相比而言,MPR(100)的公平性较好。

上述实证说明,理论上能较好解决学科差异的期刊学术影响力比较指标RIF和MPR,相对于IF这类没有考虑不同学科比较的评价指标,在“公平性测试”方面表现要好,但是就自身而言,离F. Radicchi和C. Castellano提出“公平性”评价指标的标准还相差甚远。从实证结果来看,不同学科比较的评价指标依然需要进一步改进和完善。

4 结 语

MPR与RIF两者方法不同,最终归一化的结果也有差异,虽然MPR在公平性测试中表现更好,但是与RIF并没有拉开较大的差距,且从整体表现来看,无论是MPR还是RIF距离其“公平性”指标的要求均相差甚远。尽管MPR在“公平性测试”中的“表现”不算完美,但是笔者对于该方法抱有非常乐观的看法。它不仅不是归一化方法,而且能解决期刊论文被引频次偏态分布影响期刊评价指标合理评价期刊学术影响力的问题。越来越多的科学计量学家对它表示认可,L. Leydesdorff和L. Bornmann在此基础上还提出integrated impact indicator^[19-20]。相信不断完善的MPR在期刊评价和科研评价中会有更好的表现。

评价指标研究中,指标的检验和对比分析是将指标应用于实践的前提和重要环节之一,同时也是难点之一。本文引入的“公平性测试”在国外属于较新的研究内容,在国内更鲜被提及,该方法不仅为不同学科期刊学术影响力评价指标在实践中的适用性提供了检测方法,也给期刊评价指标的检验研究带来了曙光。

参考文献:

[1] 党亚茹. 学科影响因子的时间因素差异分析[J]. 情报学报, 2002, 21(3): 328-333.

- [2] 杨兵, 彭超群, 袁赛前, 等. 相对影响因子在期刊评价和科研评价中的应用[J]. 编辑学报, 2008, 20(5): 466-468.
- [3] 任胜利, 王庆宝, 郭志明. 应慎重使用期刊的影响因子评价科研成果[J]. 科学通报, 2000, 45(2): 218-221.
- [4] 邱均平, 李爱群. 国内外期刊评价的比较研究[J]. 重庆大学学报(社会科学版), 2007, 13(3): 60-65.
- [5] 袁培国, 吴向东, 马晓军. 多指标加权值对多学科学术期刊排序初探[J]. 重庆大学学报(社会科学版), 2009, 15(4): 74-82.
- [6] 贾贤, 黄冬华. 学术期刊评价中存在的问题[J]. 北京科技大学学报(社会科学版), 2011, 27(2): 39-43.
- [7] 杜志波, 汤先忻. 标准化影响因子在不同学科期刊比较中的应用[J]. 编辑学报, 2007, 19(2): 158-160.
- [8] Radicchi F, Castellano C. Testing the fairness of citation indicators for comparison across scientific domains: The case of fractional citation counts[J]. Journal of Informetrics, 2012, 6(1): 121-130.
- [9] Garfield E, Merton R K. Citation indexing: Its theory and application in science, technology, and humanities[M]. New York: Wiley, 1979.
- [10] Moed H F. The source-normalized impact per paper (SNIP) is a valid and sophisticated indicator of journal citation impact[OL]. [2013-11-20]. <http://arxiv.org/abs/1005.4906>.
- [11] Leydesdorff L, Opthof T. Normalization at the field level: Fractional counting of citations[OL]. [2013-11-20]. <http://arxiv.org/abs/1006.2896>.
- [12] Waltman L, Van Eck N J. A general source normalized approach to bibliometric research performance assessment[OL]. [2013-11-20]. <http://www.ludowaltman.nl/doc/STI2010.pdf>.
- [13] Herranz N, Ruiz-Castillo J. Sub-field normalization in the multiplicative case: verage-based citation indicators[J]. Journal of Informetrics, 2012, 6(4): 543-556.
- [14] Colliander C, Ahlgren P. The effects and their stability of field normalization baseline on relative performance with respect to citation impact: A case study of 20 natural science departments[J]. Journal of Informetrics, 2011, 5(1): 101-113.
- [15] Zitt M. Citing-side normalization of journal impact: A robust variant of the audience factor[J]. Journal of Informetrics, 2010, 4(3): 392-406.
- [16] 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.
- [17] Moed H F. Measuring contextual citation impact of scientific journals[J]. Journal of Informetrics, 2010, 4(3): 265-277.
- [18] Moed H F. CWTS crown indicator measures citation impact of a research group's publication oeuvre[OL]. [2013-11-20]. <http://arxiv.org/abs/1003.5884>.

(下转第94页)

定量比较各位教授的 h 指数、g 指数、e 指数和 I_f 进一步证明: g 指数对引频非常敏感,但 h 指数不敏感; g 指数包含了 h 指数和 e 指数的主要特性,从文献计量学的角度看,如果利用第一作者或通讯作者论文的引频数据计算 g 指数, g 指数是一个较好的学术水平的评价指标。

参考文献:

- [1] Hirsch J E. An index to quantify an individual's research output [J]. Proceedings of National Academy of Sciences 2005, 102(46): 16569 - 16572.
- [2] Bartneck C, Kokkermans S. Detecting h-index manipulation through self-citation analysis[J]. Scientometrics 2011, 87(1): 85 - 98.
- [3] Hirsch J E. An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship [J]. Scientometrics 2010, 85(3): 741 - 754.
- [4] 马妍, 何苗, 邢星, 等. H 指数与类 H 指数应用于人才遴选的可行性探讨[J]. 情报科学 2013, 31(6): 60 - 66.
- [5] 刘合艳, 房俊民, 苑彬成. h 指数研究及应用概述[J]. 情报理论与实践 2009, 32(11): 1 - 5.
- [6] 张垒, 唐恒. 影响 h 指数、g 指数、影响因子因素的相关性研究[J]. 图书情报工作 2009, 53(20): 139 - 143.
- [7] Egghe L. An improvement of the h-index: The g-index [J]. ISSI Newsletter 2006, 2(1): 8 - 9.
- [8] Zhang Chungting. The e-index, complementing the h-index for excess citations[J]. PLoS ONE, 2009, 4(5): 1 - 4.
- [9] Egghe L. Theory and practice of the g-index [J]. Scientometrics, 2006, 69: 131 - 135.

Relations Among g, h and e Indexes and Its Bibliometrics Meaning

Sui Guiling

Library of Jilin University, Changchun 130012

[Abstract] Based on the definition of h, g and e indexes, an inequality and a formula of the functional relation among them are established by mathematical derivations and confirmed by data of literature search. By statistical analysis of the data of literature search, a mathematical equation regarding cited times as a function of sort order of papers is given when they are arranged in the descending order of their number of citations. By comparison of the indexes calculated with the data of the literature search, it is confirmed that the g is very sensitive to the number of citations, and the origin of such sensitivity is also analyzed. A discussion on how to use the g index to evaluate academic level of a scholar is also presented.

[Keywords] h-index g-index e-index bibliometrics cited times

(上接第 89 页)

- [19] Leydesdorff L, Bornmann L. Percentile ranks and the integrated impact indicator (I3) [J]. [2013 - 11 - 20]. <http://arxiv.org/abs/1112.6281>.
- [20] Leydesdorff L, Radicchi F, Bornmann L, et al. Field-normalized impact factors: A comparison of rescaling versus fractionally counted IFs [OL]. [2013 - 11 - 20]. <http://arxiv.org/abs/1211.2571>.
- [21] Leydesdorff L, Bornmann L. Integrated impact indicators compared with impact factors: An alternative research design with policy implications [J]. Journal of the American Society for Information Science and Technology, 2011, 62(11): 2133 - 2146.

Empirical Study on Methods of Comparing the Academic Impact of Journals of Different Disciplines

Chen Fuyou^{1,2} Yang Liying¹ Ding Jielan¹

¹National Science Library, Chinese Academy of Sciences, Beijing 100080

²University of Chinese Academy of Sciences, Beijing 100190

[Abstract] This paper describes the common issues about comparing the academic impact of journals of different disciplines and analyzes their origin. Then it makes a systematical classification of the existing solutions. Finally, it figures out which is better of relative impact factor (RIF) and mean percentile rank (MPR) to evaluate the academic impact using the fairness test method.

[Keywords] journal academic impact disciplinary normalization relative impact factor mean percentile rank the fairness test