

中国知网收录的国家自然科学基金论文的被引情况分析

董建军^{1,2)}

收稿日期:2012-02-27

修回日期:2012-06-28

1) 新乡医学院期刊社《眼科新进展》编辑部,453003 河南新乡,E-mail:djj1919@163.com

2) 河南省科技期刊研究中心,453003 河南省新乡市新乡医学院

摘要 本文统计分析了中国知网2001年至2010年学术网络出版总库所收录的国家自然科学基金论文的被引分布情况和历年零被引率的波动情况。研究发现国家自然科学基金论文的被引分布呈现高被引和低被引的双峰分布。国家自然科学基金论文发表后的前4年为其零被引转化成非零被引的关键时段。2010年发表的国家自然科学基金论文两年后的零被引率达到了52%,数据库中2010年发表的该类论文中有51461篇论文截至2012年2月未获得引用。因此我们认为科技期刊编辑工作者应严把国家自然科学基金论文的监督和审查,使其真正实至名归;由于受到基金论文被引双峰分布和高零被引率的影响,期刊评价指标基金论文比与期刊的影响因子相关性不大;同时国家自然科学基金论文的零被引变化规律也提示2年影响因子评价期刊存在局限性。

关键词 国家自然科学基金 影响因子 被引频次 基金论文比 零被引

国家自然科学基金是推动我国自然科学基础研究发展和促进基础学科建设,培养优秀科技人才的一种重要基金^[1]。国家对于国家自然科学基金的经费投入也是很高的,按照2010年经费概况的报道,国家自然科学基金资助项目的资金总投入已经到达99.3亿元^[2]。许多期刊的编辑和研究人员大多认为国家自然科学基金所资助的论文质量一定很高^[3]。我们前期研究发现,截至2012年2月中国知网收录的国家自然科学基金所资助的基金论文数占有基金论文总数的37.9%,然而此类基金论文总的篇均被引频次仅为5.57,比人们想象的要低很多。刘雪立和俞立平等学者对于基金论文比的研究发现基金论文比与期刊的影响因子相关性不大^[4-5]。这些现象说明人们对于基金论文的质量存在矛盾的认识。我们设计了本次研究课题,通过分析国家自然科学基金论文的被引分布和零被引情况来对国家自然科学基金论文的质量以及被引分布规律做以分析。

1 研究对象和数据的收集与分析

1.1 研究对象和数据收集

本研究的研究对象为中国知网学术文献总库的中国学术期刊网络出版总库所收录的国家自然科学基金论文。分别检索出2001年至2010年10年间中国学术期刊出版总库中所收录的国家自然科学基金论文的各频次的被引情况,收

集数据库历年发表论文截至到2012年2月的零被引文献数,计算出零被引率。

1.2 数据的分析

分析10年间发表的国家自然科学基金论文各频次被引用的分布情况。通过样本均数与总体均数的统计学推断方法,比较历年国家自然科学基金论文和数据库所有论文的零被引率,以此反向评价指标来分析国家自然科学基金论文的质量,同时分析数据库中论文的被引分布规律。

2 结果与分析

2.1 2001年至2010年发表的国家自然科学基金论文的被引分布情况

2001年至2010年发表的国家自然科学基金论文各频次被引情况见表1。从表1中可以看出,10年来国家自然科学基金论文的数量是不断上升的,然而其零被引论文的数量也呈急剧的上升。2010年其零被引论文数占到了同年该类论文总数的52%,虽然时间因素也会影响论文的被引用,但是从数据分析来看10年中零被引率最低的2003年其零被引率也达到了7%,10次以上被引率最高的也是2003年的38.5%,最少的为2010年的0.7%。国家自然科学基金论文的引用分布呈现一种偏态性,高被引和低被引大体呈现双峰分布的特点,而并非正态分布。2007年到2010年间其零被

引的论文数高于被引 10 次以上的论文数。2001 年至 2006 年后者远高于前者。2001 年至 2006 年发表的该类论文的被

引分布基本达到了稳定的状态,说明国家自然科学基金论文发表后 6 年论文的被引用达到较稳定状态。

表 1 2001 年至 2010 年发表的国家自然科学基金论文各频次被引情况

(篇)

发表年份	文献总数	被引次数											
		0	1	2	3	4	5	6	7	8	9	10	10 以上
2001	31391	3089	2689	2367	2004	1857	1623	1378	1341	1139	990	916	11998
2002	35912	3402	3001	2711	2327	2117	1843	1711	1449	1349	1219	1090	13693
2003	41727	2783	3609	3270	2896	2652	2281	2068	1813	1677	1359	1274	16045
2004	50576	4961	4669	4321	3699	3173	2940	2574	2293	2009	1768	1565	16604
2005	61460	6793	6476	5663	5138	4340	3699	3202	2887	2512	2190	1840	16720
2006	74208	9514	9049	8164	6630	5677	4784	4078	3394	2801	2455	2094	15568
2007	83859	13781	12570	10413	8466	6837	5395	4226	3547	2884	2333	1897	11510
2008	90610	19040	16899	12960	9773	7215	5354	3931	3176	2207	1910	1543	6602
2009	95752	29484	21374	12416	9615	6251	4112	2829	1959	1362	995	718	4637
2010	98364	51461	23742	10952	5350	2758	1467	839	509	301	192	103	690

2.2 2001 至 2010 年发表的国家自然科学基金论文零被引率与总体零被引率的比较

2001 年至 2010 年国家自然科学基金论文零被引率与总体零被引率的比较见表 2。通过比较我们发现历年基金论文的零被引率均远远低于总体的零被引率。历年来所有论文的零被引率基本上都在 50% 左右,而 2009 年至 2010 年超过了 60% 以上。对于基金论文来说仅有 2010 年的零被引率(52%) 高于 50%,其他均较低。2001 年至 2010 年国家自然科学基金论文零被引率与总体零被引率变化情况见图 1。图 1 中显示,国家自然科学基金论文各年的零被引率在一个较

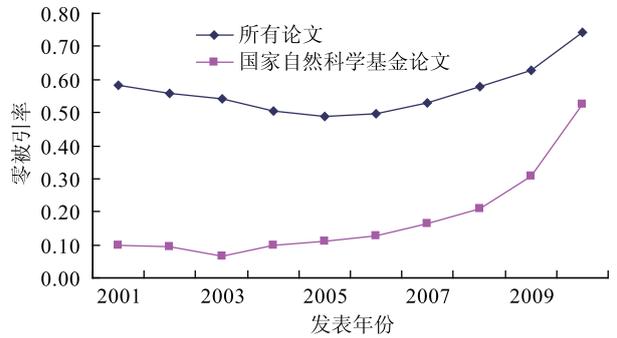


图 1 2001 年至 2010 年发表的国家自然科学基金论文零被引率与数据库所有论文零被引率变化情况

表 2 2001 年至 2010 年国家自然科学基金论文零被引率与总体零被引率的比较

年份	数据库所有论文			国家自然科学基金论文			u 值	P 值
	文献总数	0 被引文献数	0 被引率	文献总数	0 被引文献数	0 被引率		
2001	1335330	778741	0.58	31391	3089	0.10	1136.22	<0.001
2002	1438009	803288	0.56	35912	3402	0.09	1120.26	<0.001
2003	1575773	852189	0.54	41727	2783	0.07	1194.29	<0.001
2004	1637619	829579	0.51	50576	4961	0.10	1045.57	<0.001
2005	1794766	872600	0.49	61460	6793	0.11	1006.93	<0.001
2006	1976229	977422	0.49	74208	9514	0.13	1030.17	<0.001
2007	2149137	1133541	0.53	83859	13781	0.16	1066.23	<0.001
2008	2346541	1358234	0.58	90610	19040	0.21	1143.86	<0.001
2009	2423239	1520706	0.63	95752	29484	0.31	1029.17	<0.001
2010	2498369	1850625	0.74	98364	51461	0.52	784.714	<0.001

低的水平波动,2007年到2010年的波动幅度较大,2001年至2006年其波动较为稳定。基金论文零被引率的下降周期较长,下降幅度较大,随着时间的推移,特别是论文发表后的前4年,很多论文的被引频次会不断增加。

3 讨论

3.1 国家自然科学基金论文的质量问题

从我们研究的结果可以看出国家自然科学基金论文虽然比一般论文的零被引率低,但是从国家自然科学基金论文自身来看,存在着很多质量较差的论文。很多论文在发表之后的3年内也很难获得较高的被引用。即使到了发表之后的5年也有10%左右的论文未被引用。这一方面体现了我国科研水平的基本面,也从侧面反映了我们期刊编辑人员没有为此而把好关。国内期刊对于国家自然科学基金论文比较追捧,而忽视了对这类论文的监管和审查。如果将国家自然科学基金资助的论文视为高质量的论文,那么发表在2010年的论文在2012年2月进行统计时还有高达52%的论文1次也未被引用,这一结果反映到期刊的影响因子上,很明显会使国内期刊的2年影响因子普遍降低。

3.2 国家自然科学基金论文与基金论文比

我们前期研究发现国家自然科学基金论文在全国基金资助类论文中占的比例较大。其被人们视为高质量论文^[6-7],但是很多学者研究发现基金论文比与期刊的影响因子相关性不大^[4-5,8]。陈留院^[9]认为人为因素导致了基金论文比评价指标的不客观性。我们本次研究发现,国家自然科学基金论文被引频次的分布情况呈现高被引和零被引的双峰分布。而且2009年至2010年的零被引论文数和零被引率都很高。这说明基金论文很难对期刊2年内的被引频次造成正面的影响。因此基金论文比很难与期刊2年影响因子产生高度的正相关性。

3.3 零被引率与期刊的影响因子

本次研究中我们发现2006年至2010年发表的国家自然科学基金论文的零被引率呈现明显的递减趋势,特别是2008年至2010年波动的递减速度较快。2005年及以前发表的国家自然科学基金论文的零被引率波动不大。这表明

国家自然科学基金论文的被引频次在其发表的前几年,特别是前4年受到时间因素的影响较大。因此我们认为期刊两年影响因子对期刊的被引用情况的反映不够准确。忽视了国内论文被引用的时间分布规律。方红玲^[10]研究期刊论文发表后的被引情况发现,我国科技期刊论文被引量的峰值年代在论文发表后的第7~8年。我们本次研究证明国内期刊国家自然科学基金论文发表的前4年,是论文由零被引转向非零被引的关键时段。对期刊质量的评价不能仅仅局限于影响因子静态指标,而应注意其4年被引情况的变化。

参考文献

- 汪云林,李丁,付允等.国家自然科学基金合作网络分析——以中国西部环境和生态科学为例.研究与发管理,2008,20(2):102-106
- 国家自然科学基金委员会.国家自然科学基金2010年度经费概况[EB/OL][2012-02-16].<http://www.nsf.gov.cn/nsfc/cen/ndbg/2010ndbg/04/index.html>,2011-01-16/2011-12-13
- 曹玲,周广西,朱紫阳.国家自然科学基金资助大气科学领域国内论文统计与研究热点分析.中国科学基金,2011,25(4):209-214
- 刘雪立,方红玲,周志新等.科技期刊反向评价指标——零被引论文率及其与其他文献计量学指标的关系.中文科技期刊数据库,2011,22(4):525-528
- 俞立平,潘云涛,武夷山.学术期刊来源指标与影响力关系的实证研究.科研管理,2010,31(6):173-179
- 张长城,林娜,田松等.国家自然科学基金中医学和中药学学科2008年度面上项目、青年基金和地区项目受理与资助情况分析.中医杂志,2011,(14):1180-1183
- 孙占峰,初鸿,郑秋鹏.东北农业大学加强国家自然科学基金资助项目管理工作的回顾.中国科学基金,2011,(02):115-117,121
- 夏朝晖.基金论文比在科技期刊评价体系中的作用探析.中国科技期刊研究,2008,19(4):574-577
- 陈留院.人为因素对科技期刊基金论文比评价指标的影响.中国科技期刊研究,2009,20(4):634-636
- 方红玲.我国科技期刊论文被引量和下载量峰值年代——多学科比较研究.中国科技期刊研究,2011,22(5):768-710