

科技期刊论文首次被引的幂律分布规律研究¹⁾

高晓培 袁军鹏 马峥 武夷山

(中国科学技术信息研究所, 北京 100038)

(发表于《情报学报》2015 年第 7 期)

摘要: 本文以 SCIE 收录 1920、1960、2000 年的 1,118,199 篇论文为数据来源, 分析了其首次被引时间间隔的分布规律。研究结果发现, 不同出版年论文首次被引时间间隔服从幂律分布。不同学科论文在不同出版年、同一出版年的首次被引比例峰值 C 与幂指数 α 均正相关。将同一出版年的论文按照被引频次高低倒序排序并分为五组, 各组首次被引时间间隔亦满足幂律分布, 幂指数 α 与该组论文的平均被引频次正相关。随后本文探讨了首次被引幂律分布规律的两种可能的应用: 可用于构建零被引比例时间演化模型和解释科技期刊论文中迟滞承认现象。

关键词: 首次被引 零被引 人类行为动力学 幂律分布 首次被引时间间隔

分类号 G350

Power Law distribution of first citations to S&T journal papers

Gao Xiaopei, Yuan Junpeng, Ma Zheng, Wu Yishan

(Institute of Scientific and Technical Information of China, Beijing 100038)

ABSTRACT: Based on the data on 1,118,199 papers published in different years (1920, 1960, 2000) drawn from SCIE, we study the distribution of time interval between the publication year and the year of first citation. The research results show that such time intervals obey a power-law distribution. For papers published in different years or papers of different subjects published in the same year, the peak of the time interval (C) for the first citation has notable positive correlation with the power exponent (α). Then papers of the same publication year are listed in total cites order and then divide into five groups. It turns out that time intervals for the first citation for those five groups also obey a power-law distribution, with the power exponent (α) having notable positive correlation with average total cites. The power law distribution of time intervals for the first citation has two possible applications: it can be used in building the time-dependent zero citation model and in explaining the phenomenon of delayed recognition.

KEYWORDS: First citation. Zero citation. Human behavior dynamics. Power law. time interval for the First citation.

CLASS NUMBER: G350

收稿日期: 2014 年 9 月 27 日。

作者简介: 高晓培, 女, 1986 年生, 硕士研究生, 研究方向: 科学计量学; 袁军鹏, 男, 博士, 研究方向: 情报学、科学计量学; 马峥, 男, 助理研究员, 研究方向: 情报学、科学计量学; 武夷山, 男, 研究员, 研究方向: 科学计量学、情报研究。

¹⁾ 基金项目: 国家自然科学基金项目资助“论文零被引的时间演化规律、影响因素及其应用研究”(编号:71373252)

1 引言

科技期刊论文是科技交流的重要形式之一,论文被阅读并被引用是其交流价值得以体现的重要表现形式。论文首次被阅读是论文在科技交流系统中发挥作用的起点,其首次被引用则是可被观测到的交流效果的最早证据。对论文首次被引问题的研究有助于我们了解论文在科技交流系统中发挥作用的情形,是文献计量学领域关注的内容之一。

目前围绕论文首次被引问题展开的研究可以概括为两个主要的方面:建立首次被引比例的分布模型^[1-4]和基于首次被引过程设计科技期刊评价指标^[5-7]。以首次被引比例分布模型为研究目标,R. Rousseau 最早用双指数模型描述了论文的首次被引过程,在该模型中,他假定每年未被引论文向被引论文转化的速率满足指数分布^[2]。此后,L. Egghe 结合洛特卡定律对该模型做出了改进^[1],使之能适用于凹形和 S 形曲线。2011 年,Quentin L. Burrell 基于非齐次泊松过程建立了首次被引的随机模型^[3],其后他继续研究了论文的第 n 次被引用的分布情况^[4]。这些都是从传统的文献计量学视角出发,对首次被引问题进行抽象并建模。

2005 年 Barabási 发现人类活动具有“阵发(burst)”和“重尾(heavy tails)”的特征,其时间分布具有高度的非均匀性,服从幂律分布,开创了人类行为动力学研究的新方向^[8]。此后,在人类行为多种案例的分析中,如水陆或电子通信^[9,10]、图书借阅^[11,12]、博客评论^[13]、QQ 群消息^[14]、在线电影点播^[15]等行为中,均发现了时间间隔的幂律分布规律,且幂指数在 1~3 之间。在文献计量学领域,幂律分布虽然被用于描述词频分布、被引频次分布^[16-18]等,但是迄今未见用幂律分布模型来拟合首次被引过程的研究。我们认为,一组论文,以其出版年(t_p)为起点,论文在不同的年份(t_c)被首次引用,则每篇论文首次被引都对应于一个首次被引时间间隔($t_c - t_p$)。对论文首次被引的时间间隔是否满足幂律分布这一问题的研究不仅可以丰富人类行为动力学的内容,还可以继双指数模型和随机模型之后,增加一种描述首次被引过程的模型,对研究首次被引以及相关的文献计量学研究具有一定的意义。

2 幂律分布模型

本文尝试建立一组论文首次被引过程的幂律分布模型。研究中几个概念的定义如下:

定义 1 首次被引的时间间隔:指首次被引年与出版年之间的时间间隔。用 t 表示。

$$t = t_c - t_p \quad (\text{公式 } 1)$$

t_p 表示该论文的出版年, t_c 表示该论文获得首次引用的时间。 t 的数值反映了该论文由“未被引”到“被引”状态转换的时间间隔^[19]。在本文的研究中,出于数据可获取性考虑,以年为单位来定义该时间间隔 t ,一篇论文如果在出版当年即被引用,则其 t 值为 0;如果在出版次年被引用,则其 t 值为 1;依次类推。

定义 2 首次被引论文数:指某年出版的一组论文中,在统计年之前从未被引用、在统计年首次被引用的论文数量。用 $S(t)$ 表示。其中, t 代表统计年与出版年的时间间隔,即首次被引时间间隔。

定义 3 首次被引论文比例:指某年出版的一组论文中,在统计年的首次被引论文数占该组论文的比例。用 $R(t)$ 表示。其中, t 代表统计年与出版年的时间间隔,即首次被引时间间隔。

借鉴人类行为动力学研究中时间间隔的幂律分布规律,我们认为,某年出版的 D 篇论

文，其首次被引论文数和首次被引时间间隔的关系服从幂律分布。这种幂律分布规律可以表示为：

$$S(t) = \begin{cases} E, & t = 0 \\ Bt^{-\alpha}, & t \in \{1,2,3 \dots\} \end{cases} \quad (\text{公式 2})$$

$S(t)$ 是首次被引时间间隔 t 所对应的首次被引论文数， E 、 B 和 α 均是常数。

根据 $S(t)$ 和 $R(t)$ 的定义知 $S(t) = D \cdot R(t)$ ，代入公式 2 可得：

$$D \cdot R(t) = \begin{cases} E, & t = 0 \\ Bt^{-\alpha}, & t \in \{1,2,3 \dots\} \end{cases} \quad (\text{公式 3})$$

公式 3 两边同时除以 D ，可得：

$$R(t) = \begin{cases} \frac{E}{D}, & t = 0 \\ \frac{B}{D}t^{-\alpha}, & t \in \{1,2,3 \dots\} \end{cases} \quad (\text{公式 4})$$

令 $M = E/D$ ， $C = B/D$ ，则公式 4 可以简写为：

$$R(t) = \begin{cases} M, & t = 0 \\ Ct^{-\alpha}, & t \in \{1,2,3 \dots\} \end{cases} \quad (\text{公式 5})$$

M 、 C 、 α 均为常数。公式 2 和公式 5 可以解读为：同一组论文，其首次被引论文数和首次被引论文比例均服从幂指数为 α 的幂律分布。 α 表示论文首次被引比例的衰减情况； M 是论文出版当年的首次被引比例；当 $t=1$ 时，首次被引比例数值为 C 。

3 数据说明及拟合结果

3.1 数据说明

为了验证首次被引幂律分布模型的合理性，本文采用了汤森路透公司 Web of Science 平台 SCIE 数据库中的数据考察了公式 5 对实际数据的拟合效果。相关过程简述如下：

（1）数据采集。按照检索式“PY=1920 OR PY=1960 OR PY=2000 数据库：SCIE”的检索结果下载数据，将数据导入数据库；通过 WOS 平台的创建引文报告功能获得每篇论文在每年的被引次数并添加至上述数据库。论文总量逾 111.8 万篇（详见表 1），数据下载时间为 2014 年 6 月。

（2）数据处理：根据后文分析的需要，在数据库中添加标记字段，进行数据分组；添加字段，标记每篇论文在出版后不同时间间隔的被引次数。

（3）数据统计与分析。在此阶段使用的工具包括 Foxpro、OriginPro 8 和 MATLAB。

表 1 样本论文集的主要文献计量学指标

出版年		1920 年	1960 年	2000 年	总计
论文数*（篇）		17410	109733	991056	1118199
篇均被引频次（次/篇）		4.77	19.17	21.04	
出版后 10 年	零被引论文数（篇）	12347	34876	288283	
	零被引论文比例（%）	70.92	31.78	29.09	
截至 2013 年	零被引论文数（篇）	9741	30546	272203	
	零被引论文比例（%）	55.95	27.84	27.47	

* 论文类型包括: Article, Proceedings Paper, Bibliography, Book Review, Correction, Database Review, Editorial Material, Hardware Review, Letter, Meeting Abstract, News Item, Poetry, Reprint, Software Review。由于 SCIE 数据库关于论文被引频次的最小时间单位是年, 本文时间间隔统计单位为年。

3.2 幂律分布模型的拟合结果

论文集 D 的首次被引论文数与首次被引论文比例是否服从幂律分布是本节考察的内容之一。同时, 由于科技期刊论文的被引情况受多种因素的综合影响 (如时间、学科、语种、文献类型等), 以不同分类方式划分的论文集合是否服从幂律分布也是本节要考察的内容。

(1) Aritcle、Review 类型论文首次被引数和首次被引比例均服从幂律分布

如表 1 所示, 本文所采集的论文共计 111.8 万篇, 其中 Article 类型的论文 71.5 万篇、Review 类型论文 3.0 万篇和其他类型论文 24.6 万。除 Article、Review 之外的 24.6 万篇论文包括 Book Review、Editorial Material 等, 这些文献虽然学术性可能不及前两种文献, 但是实际分析显示这些论文也在学术交流中发挥了重要作用, 这种作用表现为在学术交流中的被引, 其篇均被引频次为 2.14 次/篇。

全部 111.8 万篇论文按照表 2 所示的结果分组, Article 类型、Review 类型和其他类型论文的首次被引论文数和首次被引比例均服从幂律分布, 幂指数 α 的值分别是: 1.45 (Article 类型)、1.85 (Review 类型) 和 1.16 (其他类型), 反映拟合效果的 R^2 值分别是 0.98、0.99 和 0.97。

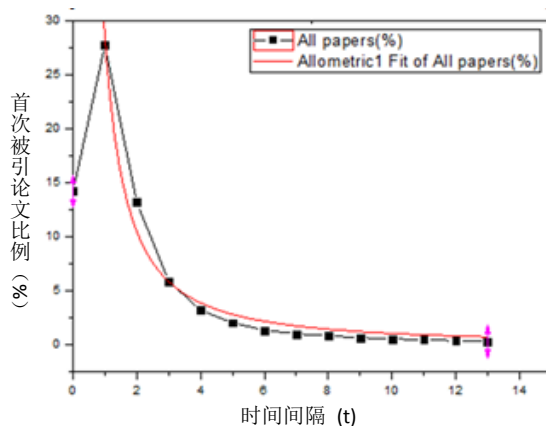
表 2 不同类型论文总数和篇均被引频次

	1920 年		1960 年		2000 年		总计	
	ACPP	TA	ACPP	TA	ACPP	TA	ACPP	TA
Article	5.62	14444	27.21	69100	25.50	714868	25.29	798412
Review	4.11	159	49.10	936	72.96	29892	71.88	30987
others	0.44	2807	4.46	39697	1.78	246296	2.14	288800

* Article 包括 SCIE 数据库标记为 “Article ; Proceedings Paper” 和 Article 类型的论文; ACPP 是篇均被引频次 (Average Citations Per Paper) 的简写, 单位: 次/篇; TA 是总论文数 (Total Articles) 的简写, 单位: 篇。

(2) 全部 111.8 万篇论文首次被引数和首次被引比例均服从幂律分布

图 1 是全部 111.8 万篇论文首次被引比例的拟合效果图, 幂指数 α 的值为 1.43, 反映拟合效果的 R^2 值为 0.98。数据拟合结果显示, 这 111.8 万篇论文其首次被引论文数亦服从 $\alpha = 1.43$ 的幂律分布。

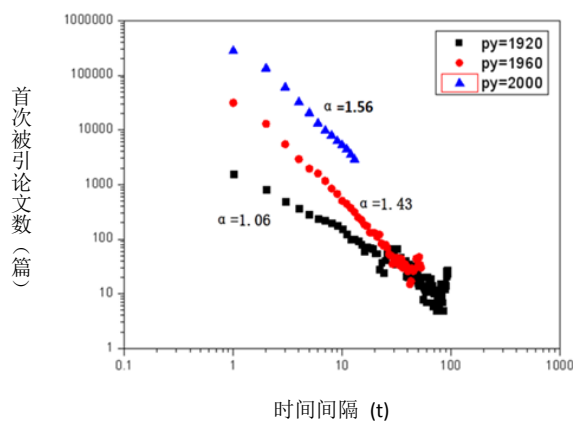


注：图中双向箭头标记的是 OriginPro 软件自动标记的拟合的起止位置。

图 1 全部 111.8 万篇论文首次被引比例拟合效果图

(3) 不同出版年论文集合首次被引论文数和首次被引比例服从幂律分布

图 2 是双对数坐标下不同出版年论文中首次被引论文数随时间间隔变化的规律。如图 2 所示，出版年为 1920 年、1960 年和 2000 年的三组论文在不同时间间隔下首次被引论文数均服从幂律分布 $S(t) \sim t^{-\alpha}$ ，其幂指数值分别为 1.06、1.43 和 1.56。对比图中的三组拟合结果，出版年为 2000 年的论文较其他两组出现了更为“干净”的尾部，这是因为该年数据点较少，随着时间范围的增大，较小的数值变化都可能引起较大的波动，造成图中数据点的堆叠。



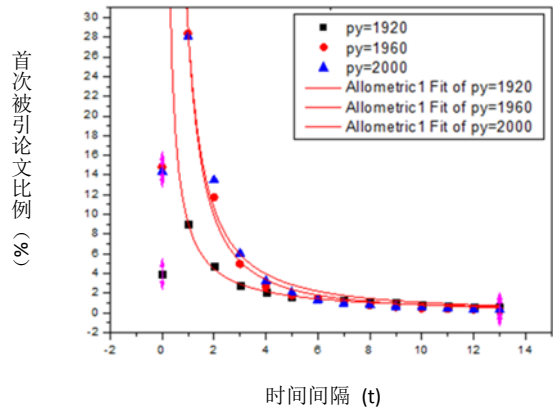
注：py = 1920 表示出版年为 1920 年，后同。不同出版年时间间隔取值范围不同，py = 1920 时 $t \in [0, 93]$ ，py = 1960 时 $t \in [0, 53]$ ，py = 2000 时 $t \in [0, 13]$ 。

图 2 不同出版年论文中首次被引论文数随时间间隔的变化

由公式 2-公式 5 的数学推导部分可知，若某一论文集合的首次被引论文数服从参数为 α 的幂律分布，则其首次被引比例亦服从参数为 α 的幂律分布，反之亦然。由于首次被引论文数是一个绝对数量，不宜用于不同规模间样本的比较，在此重点讨论首次被引比例的幂律分布规律。同时，为了消除如图 2 中所示数据点数量对拟合趋势的影响，后文首次被引比例的拟合中，三组不同出版年的数据都只取前 14 个数据点，拟合效果见图 3，相关参数见表 3，

三组数据的首次被引比例均能很好地拟合幂律分布，反映数据拟合效果的 R^2 这一指标均大于等于 0.98。当拟合的数据点数量变化时，幂律参数的大小会受到影响，但变化幅度较小。如表 3、图 2 对比所示，当拟合数据点减少后，幂律参数出现了微小的变化，由出版年为 1920 年的论文由 1.06 变成了 1.03；出版年为 1960 年的论文由 1.56 变成了 1.55。

此外，我们对 2000 年的数据采取了不同的分组方式对数据进行拟合，如按照学科、语种、文献类型（论文、综述等）进行分组，各组数据均服从幂律分布。受限于篇幅，这里不做介绍。



注：图中双向箭头标记的是 OriginPro 软件自动标记的拟合的起止位置。

图 3 不同出版年论文中首次被引比例随时间间隔的变化

表 3 不同出版年 $t=0\sim13$ 时数据拟合参数

出版年	拟合参数			首次被引比例（%，统计值）	
	C	α	R^2	t=0	t=1
1920 年	9.13	1.03	0.9975	3.93	9.06
1960 年	28.74	1.55	0.9920	14.83	28.37
2000 年	28.68	1.43	0.9816	14.33	28.02

4 被引分布规律在零被引问题研究中的应用

4.1 零被引时间演化模型

构建论文零被引时间演化模型是零被引问题研究的内容之一。目前已有一些学者尝试建立零被引比例随时间变化的数学模型，如，Burrell^[20]提出了一个未被引论文比例随时间变化的模型 $P(X_t = 0) = \frac{a}{(v-1)t} [1 - (\frac{a}{a+t})^{v-1}]$ ，但该模型与真实数据的拟合效果并不理想。胡泽文等以一个三参数负指数函数 $P(X_t = 0) = K + Ae^{-t/s}$ 拟合零被引的时间演化规律并以 6 种期刊为例进行了实证研究，并获得了较好的拟合效果^[21]。上述模型是直接以论文零被引比例为变量进行参数拟合，寻找论文零被引比例的变化规律。由于论文“被引”与“未被引”是一对互补的概念，因此，科技期刊论文首次被引幂律分布规律的应用价值之一在于，借助引入首次被引比例这一概念，可以建立论文集合零被引比例时间演化的模型。

如果用 $Z(t)$ 表示不同时间间隔下一组论文零被引的比例， $R(t)$ 表示不同时间间隔下一组

论文的首次被引比例，则，

$$Z(t) = 1 - \sum R(t) \tag{公式 6}$$

将公式 5 代入公式 6，可得：

$$Z(t) = \begin{cases} 1 - M, & t = 0 \\ 1 - M - \sum_{n=1}^t Cn^{-\alpha}, & t \in \{1,2,3 \dots\} \end{cases} \tag{公式 7}$$

公式 7 即一组论文零被引时间演化模型。在已知 M、C 和 α 这三个参数的情况下，可以通过此模型预测一组论文出版后每一年的零被引比例。本文以 1960 年数据为例，比较了胡泽文的三参数负指数模型和公式 5 模型的拟合效果，事实证明，两个模型都能获得了较好的拟合效果，本文的模型在两组数据（1960 年、2000 年）中获得了更好的拟合优度，拟合结果见表 4。

表 4 零被引三参数负指数模型和本文模型拟合效果对比

模型	数据	参数	R^2
三参数负指数模型 ^[21] : $P(X_t = 0) = K + Ae^{-t/s}$	1920 年	K=57.64、A=29.14、s=14.30	0.9570
	1960 年	K=29.02、A=53.78、s=1.88	0.9752
	2000 年	K=29.94、A=56.04、s=1.60	0.9951
本文公式 7 模型: $Z(t) = \begin{cases} 1 - M, & t = 0 \\ 1 - M - \sum_{n=1}^t Cn^{-\alpha}, & t \in \{1,2,3 \dots\} \end{cases}$	1920 年	M=3.04、C=10.30、 α =1.13	0.9977
	1960 年	M=13.95、C=32.19、 α =1.81	0.9964
	2000 年	M=13.66、C=31.68、 α =1.64	0.9946

4.2 用幂律分布规律解释迟滞承认（睡美人）现象的存在

迟滞承认描述的是论文出版后一段时间内未得到认可（以被引用为标志），较晚才开始被引用，且最终被引频次较高的现象。迟滞承认的论文被比喻为科学中的睡美人（Sleeping Beauties in Sciences）^[22]。A. van Raan 用沉睡时长、沉睡深度和唤醒强度界定了睡美人的 3 个指标^[22]。在不同的研究中，虽然大家对睡美人采用了不同的界定标准，但首次被引时间与出版时间相比的延迟与较高的被引频次是迟滞承认论文最典型的两个特征。

首次被引的幂律分布规律可以用于解释迟滞承认现象。数据拟合的结果显示，按照被引频次对数据分层，各组论文的首次被引比例均符合幂律分布。

（1）分组方式

相同出版年的论文按照截至 2013 年的总被引频次从高到低降序排列，然后等分。需要指出的是，为了保证相同被引频次的论文被分入一组，每组论文数并非严格相等而是按照实际被引情况作了微调。例如，1920 年出版的论文有 17410 篇，其中 7669 篇截至 2013 年获得了首次被引，这 7669 篇论文如果等分为五组，每组论文数为 1533.8 篇。事实上，这 7669 篇论文被引频次为 1 次的论文有 1833 篇，因此第五组论文数为 1833 篇而不是 1533.8 篇；其余组亦同。之所以将所有论文分为 5 组，而不是 3 组或者 7 组，是因为分组太少不易于比较组之间的差异，分组太多并无必要，分组结果详见表 5。

（2）拟合结果

如表 5 所示，各出版年被引频次较高（第一组和第二组）、被引频次居中（第三组）和

被频次较低（第四组和第五组）的论文均服从幂律分布。这说明幂律分布在不同被引频次论文群体中分布的适用性。三种不同出版年的各五组论文，其篇均被引频次和参数 α 均正相关，相关系数分别为 0.83（1920 年）、0.90（1960 年）和 0.94（2000 年）。从第一组到第五组，随着篇均被引频次的降低，首次被引比例、参数 α 和反应拟合效果的 R^2 均出现了逐渐降低的趋势。被引频次居中的第三组论文，其参数 C、 α 在五组数据中亦居中。

（3）结果解读

由于不同被引频次分层抽样的论文集合其首次被引均服从幂律分布，其中较高频次被引论文的首次被引符合幂律分布可用于解释睡美人文献的存在。较高被引频次组别论文服从幂律分布，其时间间隔较大时首次被引的论文即对应于被迟滞承认的论文。从幂律分布的特性角度可以解释为：幂律分布的典型特征在于其不均衡性，存在一个巨大的头部和长长的尾部，巨大头部的存在，说明多数具有潜在高被引特性的论文在发表之初即可以被识别，而其长长尾部则是睡美人文献。

出版年被引频次较低的两组论文服从幂律分布可以解释为，人们在引用时更倾向于引用较新的论文，一些最终看来并不十分优秀的论文在出版的早期可能获得被引，但随着时间的推移，它们继续被引用的概率会逐渐降低。

表 5 不同出版年不同组首次被引比例拟合参数

	出版年/分组	论文数 (篇)	篇均被引频次 (次/篇)	C	α	R^2
被引频次 较高	1920 年/第一组	1369	42.22	32.77	1.42	1
	1920 年/第二组	1304	9.47	27.65	1.21	0.99
	1960 年/第一组	15599	95.22	49.14	2.43	1
	1960 年/第二组	16280	22.33	48.76	1.92	0.99
	2000 年/第一组	144739	98.61	49.46	2.93	1
	2000 年/第二组	143764	25.88	52.25	1.94	0.98
被引频次 居中	1920 年/第三组	1390	4.86	23.52	1.09	0.99
	1960 年/第三组	16414	10.04	43.76	1.58	0.99
	2000 年/第三组	142510	12.55	45.04	1.45	0.94
被引频次 较低	1920 年/第四组	1773	2.42	16.7	0.91	0.97
	1920 年/第五组	1833	1	10.22	0.72	0.94
	1960 年/第四组	15932	4.3	35.14	1.28	0.97
	1960 年/第五组	14962	1.42	23.87	1.02	0.95
	2000 年/第四组	136404	5.8	34.86	1.09	0.87
	2000 年/第五组	151436	1.82	21.06	0.68	0.85

5 结论与讨论

通过上文的分析，可以得到以下结论：

（1）不同出版年论文集合的首次被引数和首次被引比例与时间间隔的关系均满足幂律分布。以年为时间间隔统计单位，首次被引比例峰值 C 与幂指数 α 正相关。同一出版年不同学科的论文，首次被引比例峰值 C 与幂指数 α 也存在正相关关系。出版年为 1920、1960

和 2000 年的论文集合对应的幂指数分别为 1.06、1.43 和 1.56，与其他人类行为动力学研究中发现的幂指数取值范围（1~3）相一致。

（2）同一出版年的论文，按照被引频次高低进行排序、分组，各组的首次被引论文数和首次被引比例亦满足幂律分布，幂指数 α 与该组论文的平均被引频次正相关。

（3）首次被引论文数和首次被引论文比例随时间间隔变化的幂律分布特征，可以应用于对零被引问题的研究。通过零被引与首次被引的关系，可以建立如上文公式 7 所示的零被引时间演化模型。通过对真实数据的拟合，我们发现公式 7 比三参数负指数模型获得了更好的拟合效果。此外，高被引论文的首次被引时间间隔的幂律分布特征，可以用于解释迟滞承认现象。

本文的研究借鉴了人类行为动力学中人类行为时间性的幂律分布模型，但是人类行为动力学研究中的时间间隔往往是一个即时反应的连续变量，而由于期刊出版周期的不同、论文阅读引用与出版之间的时滞等原因，论文首次被引的时间间隔往往更加复杂，在此情况下幂律分布模型是否应做相应修正有待进一步的探讨。

人类行为动力学研究虽然用基于优选权决策的排队模型和基于兴趣的非齐次泊松过程模型解释了幂律分布的动力学机制，但幂指数数值上的差别能否作为区分不同类别人类行为的标尺尚未有定论^[10]。同样，无论是按照学科、文献类型、语种等维度对论文首次被引情况进行拟合，其分布均服从幂律分布，但得到的幂指数的差异是否可以反映论文在不同维度分布上的差异、造成这些差异的内在原因又有哪些，这些也都有待进一步的研究。

参考文献:

- [1] Egghe, L. A heuristic study of the first-citation distribution[J]. *Scientometrics*. 2000,48(3): 345-359.
- [2] R. Rousseau. Double exponential models for first-citation processes[J].*Scientometrics*. 1994,30(1):213-227.
- [3] Burrell, QL. Stochastic modelling of the first-citation distribution[J].*Scientometrics*. 2011,52(1): 3-12.
- [4] Burrell, QL. The nth-citation distribution and obsolescence[J]. *Scientometrics*. 2002,53(3):309-323.
- [5] A. Schubert, W. Glänzel. Mean response time — A new indicator of journal citation speed with application to physics journals[J].*Czechoslovak Journal of Physics B*.1986,36(1):121-125.
- [6] Glanzel, Wolfgang; Rousseau, Ronald; Zhang, Lin . A Visual Representation of Relative First-Citation Times[J]. *Journal of the American society for information science and technology*, 2012,63(7): 1420-1425.
- [7] Lutz Bornmann; Hans-Dieter Daniel .The citation speed index: A useful bibliometric indicator to add to the h index[J]. *Journal of informetrics*, 2010,4(3):444-446
- [8] Albert-László Barabási. The origin of bursts and heavy tails in human dynamics[J]. *Nature*, 2005,435(7039): 207-211.
- [9] 李楠楠,周涛,张宁等. 人类动力学基本概念与实证分析[J]. *复杂系统与复杂性科学*, 2008,5(2):15-24.
- [10] 樊超,郭进利,韩筱璞等. 人类行为动力学研究综述[J]. *复杂系统与复杂性科学*, 2011,08(2):1-17.
- [11] 樊超,郭进利,纪雅莉等. 基于图书借阅的人类行为标度律分析[J].*图书情报工作*, 2010,54(15):35-39.
- [12] 樊超.从图书借阅看人类群体和个体行为的动力学机制[D]. 上海理工大学, 2010.
- [13] 郭进利.博客评论的人类行为动力学实证研究和建模[J]. *计算机应用研究*,2011,28(4):1422-1424,1433.
- [14] 罗芳,杨建梅,李志宏等. QQ 群消息中的人类行为动力学研究[J]. *华南理工大学学报: 社会科学版*, 2011,13(4):14-19.
- [15] 周涛.在线电影点播中的人类动力学模式[J]. *复杂系统与复杂性科学*,2008,5(1):1-5.
- [16] S. Redner. How popular is your paper? An empirical study of the citation distribution[J].*The European*

Physical Journal B - Condensed Matter and Complex Systems. 1998,4(2):131-134

- [17] 方爱丽,高齐圣,张嗣瀛等. 引文网络的幂律分布检验研究[J]. 统计与决策,2007,(14):22-24.
- [18] 毛国敏,蒋知瑞,任蕾等. 期刊论文被引频次的幂律分布研究[J]. 中国科技期刊研究, 2014,25(2):293-298,307.
- [19] Leo Egghe, Lutz Bornmann, Raf Guns. A proposal for a First-Citation-Speed-Index[J]. Journal of Informetrics. 2011(5):181-186.
- [20] Burrell, QL. Alternative thoughts on uncitedness. Journal of the American Society for Information Science and Technology, 2012,63(7), 1466-1470.
- [21] Zewen Hu, Yishan Wu. Regularity in the time-dependent distribution of the percentage of never-cited papers: An empirical pilot study based on the six journals[J]. Journal of Informetrics, 2014(8):136-146.
- [22] A. van Raan. Sleeping Beauties in Science[J]. Scientometrics, 2004,59(3):467-472.

(责任编辑: 赵康)