ISSN: 0975-833X

Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 11, Issue, 05, pp.3712-3716, May, 2019

DOI: https://doi.org/10.24941/ijcr.35094.05.2019

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

RESEARCH ARTICLE

AN EXAMINATION OF LOTKA'S LAW OF AUTHORSHIP PRODUCTIVITY IN THE FIELD OF NEUROSCIENCES

[1]*Dr. Gavisiddappa Anandhalli and [2]Latha

[1]Associate Professor and University Librarian, Karnataka State Akkamahadevi, Women's University, Vijyapura, Karnataka, India

[2]Research Scholar, Karnataka State Akkamahadevi Women's University, Vijyapura, Karnataka India

ARTICLE INFO

Article History:
Received 24th February, 2019
Received in revised form
04th March, 2019
Accepted 05th April, 2019
Published online 30th May, 2019

Key Words:

Lotka's Law, Authorship Productivity, Authorship Pattern. Scientometrics,

ABSTRACT

Lotkas law of Scientific Productivity is Bibliometric study, it has been used to calculate the number of authors against the number of publications contributed by the authors was plotted on a logarithmic scale. The points were closely distributed around the straight line having a slope of almost 2. The present study tests the conformity of Lotka's law for authorship distribution in the field of neuroscience using the web of science core collection database. There are 57640 publications contributed by 132481 authors. Lotka's law was tested by with help methods given by Pao and Maximum Likelihood method as reported by Nicholls. Later the data was organized according to authorship pattern with the help of MS Excel. The value of c and exponent n is calculated with help of collie and Pao formula. It is observed that the number of authors publishing n papers was $1/n^2$ of those publishing one paper. The present study is almost following the same pattern. Hence, it can be concluded that Lotka's law of Author productivity is well fitted/followed in the field of Neuroscience.

**Corresponding author:*
Dr. Gavisiddappa Anandhalli

***Citation: Dr. Gavisiddappa Anandhalli and Latha,* 2019.** "An Examination of Lotka's Law of Authorship Productivity in the field of Neurosciences", *International Journal of Current Research*, 11, (05), 3712-3716.

INTRODUCTION

Bibliometrics is a potential area of research in the field of library and information science (LIS) pragmatic applications in measuring the growth and development of literature. It helps in formulating need-based collection building policy and provides authentic data to inform managers to take the judicious decision in the process of document selection. Bibliometrics is also having the wide range of applications the other field. Many research fields use bibliometric methods to explore the impact of their field, the impact of the researchers, institutions, and the impact of the paper also. The term Bibliometrics was coined by Allan Pritchard in a paper published in 1969 entitled "Statistical Bibliography or Bibliometrics". He defined the term as "the applications of Mathematics and Statistical Methods to books and other media of communication (Allan Pritchard 1969).

About Lotka's Law: Lotka's Inverse Square Law of Scientific productivity is derived by Alfred J. Lotka's and it is the first Scientometric or Bibliometric law. to obtain the data for deriving his law, Lotka's (1926) made a count of the number of personal names in the 1907 -1916 Decennial Index of Chemical Abstracts against which there appeared 1, 2, 3, etc. entries, covering only the letters A and B of the alphabet. He also applied a similar process to the name index in Felix Auerbach's Geschichtstafeln der Physic (Leipzig: J. A. Barth, 1910), which dealt with the entire range of the history of physics through 1900 by using the latter source, Lotka's hoped to take into account not only the volume of production but also quality, since it is listed only the outstanding contribution in physics. In making these counts Lotka's credited only the senior author in joint publications, on the basis of this data, Lotka's derived what he termed as Inverse Square Law According to which of 60% of authors produce one paper, 25% of authors produce two papers, 11% produce three papers and 6.3% produce four papers etc.

In the generalized form Lotka's law referred as inverse power law by Bookstien[3] and expressed as:

$$Y_x = \frac{C}{x^n};$$

Where x= 1,2,3,…, ,

Y_x represents the probability of an author to publish x times, C and n are the constant to determine for each data set;

n is constant, it is considered as an exponent which can be calculated with help of x,y, and c for the given data set of a subject. Where n value may lie between -1.2 and -3.8 for Lotka's law to be considered as Inverse Square Law.

Review of Literature: There are significant numbers of research conducted on the verification of Lotka's law with help of empirical data set. Almost many studies have confirmed the validity of a law. It is observed that the exponent value 'n' is not always 2 rather it is a variable value, the some of the important studies undertaken in this field as mentioned below. Pao has studied a Least-Squares Method to examine Lotka's Law. She suggested the procedures to calculate the values of the exponents' n and constant c and the subsequent Kolmogorov-Smirnov (K-S) Goodness-of-fit test for conformity. Further, Some weaknesses of Pao's Methodology are reported in the literature such as the fact that the Least-Squares Approach gives acceptable results only if author data are truncated.

Nicholls applied Lotka's law using authors and the Maximum Likelihood (ML) approach to estimate the parameters. In this study, it is found that the ML method is generally better. Newman observed that the Maximum Likelihood is a good method and that there is a tendency for Least-Square Fits to overestimate the scope of the power law since the statistical fluctuations in the logarithms of the data are greater in the downward direction than in the upward one. Following Nicholls's methodology, Rousseau and Rousseau developed a straightforward computer program called Lotka for determining the best fitting parameters for Lotka distribution. The program also applies a Kolmogorov-Smirnov (K-S) test for conformity.

Objectives of the study

- To test the validity of the Lotka's law in the field of Neuroscience.
- To examine K-S statistics for the applicability of the results obtained by these methods

Hypothesis

Following the null hypothesis has been formulated

H_0: The observed authorship data distribution is the same as the theoretical authorship distribution in the field of Neuroscience.

METHODOLOGY

In the present study, data is downloaded from the Web of Science Online version from 2000 to 2018 using the keyword Neuroscience. There were 57640 articles contributed by 132481 authors published in more than five thousand journals are used as the main source of data for the present study. The data has been transferred on excel sheet and calculated number of papers published by one author, two authors, three authors, and four authors so and so forth by manually by using filter mechanism. Lotka's law was tested by using various methods suggested by Pao and Maximum Likelihood method advocated by Nicholls.

Calculation of exponent 'n': The value of exponent n is calculated by the Least Square Method described by Pao (1985) using the following formula

$$N=\frac{N\sum XY-\sum X\sum Y}{N\sum X^2-(\sum X)^2} \qquad (1)$$

X is the logarithm of x, i.e. the number of articles published (x=1,2,3,4) and

Y is the logarithm of y i.e. number of authors.
N is the number of pairs of data.

To test the applicability of Lotaks law on the set of data Kolmogorov Simonov (KS) statistic suggested by Cooile is used to determine the maximum deviation D.

$$D=Max|\ F(x)-S(x)| \qquad (2)$$

Where F(x) is the theoretical cumulative frequency distribution and S(x) is the observed cumulative frequency distribution of a sample. The maximum difference is obtained and then compare with critical value to accept or reject the Null Hypothesis.

Procedure for calculation of C

The value of constant c is calculated using the following formula:

$$C=\frac{1}{\sum_1^{p-1}\frac{1}{x^n}+\frac{1}{(n-1)(p^{(n-1)})}+\frac{1}{2p^n}+\frac{n}{20(p-1)^{(n+1)}}} \qquad (3)$$

With x = 1,2,3….19

In the present data set, the value of 'c' can be calculated with the help of equation (1). To calculate first part of the above equation for the constant 'c' the value of n =2.02752 is substituted in the equation (3) by Least Square Method (value of p is chosen as 20) for that point equation givens negligible value and therefore of no use. The first part of the equation value can be calculated as $\sum_1^{p-1}\frac{1}{x^n}$ = 1.564698, same way the second part of the equation $=\frac{1}{(n-1)(p^{(n-1)})}$=.0.1298, the third part of the equation is $1/2p^n$ 0.0088 and the remaining fourth part = $n/24(p-1)^{n+1}$ = .000000, then adding of all four part and dividing them by 1 gives c=.630327

$$C=\frac{1}{1.564\ 698+0.01298+0.0088+0.0000}$$

C=.630327

Application of lotka's law: Lotka's law of Scientific Productivity is Bibliometric study, which is used to calculate the number of authors against the number of publications made by the authors is plotted on a logarithmic scale. The points are closely distributed around the straight line having a slope of almost 2. The procedure is to apply Lotka's requires a Log-log graph between the number of authors and number of papers only those pairs of points which lies on relatively straight line was considered for the calculating the parameters of the Lotka's law (Fig-1 and Table -1). Table-1 shows the frequency distribution of author productivity in the field of neuroscience research, out of 132481 authors, as many as more than 36803 authors contributed only one paper, which amounts to 63.84%

Table1. Frequency of Distribution of research output of Neurosciences

No of Articles	Number of Authors Observed	Percentage of Authors	Total Contribution
1	36803	63.84976	36803
	9074	15.74254	18148
2			
3	3761	6.524983	11283
4	2440	4.233171	9760
5	1007	1.747051	5035
6	664	1.151978	3984
7	652	1.131159	4564
8	474	0.822346	3792
9	414	0.718251	3726
10	355	0.615892	3550
11	311	0.539556	3421
12	236	0.409438	2832
13	236	0.409438	3068
14	230	0.399028	3220
15	177	0.307078	2655
16	117	0.202984	1872
17	118	0.204719	2006
18	59	0.102359	1062
19	57	0.09889	1083
20	52	0.090215	1040
21	51	0.08848	1071
22	51	0.08848	1122
23	110	0.19084	2530
24	101	0.175226	2424
25	60	0.104094	1500
31	30	0.052047	930
Total	57640	100.00	132481

Table 2. The Calculation of n exponent

No of papers(x)	Number of authors(y)	X=Log(x)	X=Log(y)	XY= X* Y	X^2=log(x)^2
1	36803	0	10.51333	0	0
2	9074	0.693147	9.113168	6.316767	0.480453
3	3761	1.098612	8.23244	9.04426	1.206949
4	2440	1.386294	7.799753	10.81275	1.921812
5	1007	1.609438	6.914731	11.12883	2.59029
6	664	1.791759	6.498282	11.64336	3.210402
7	652	1.94591	6.480045	12.60958	3.786566
8	474	2.079442	6.161207	12.81187	4.324077
9	414	2.197225	6.025866	13.24018	4.827796
10	355	2.302585	5.872118	13.52105	5.301898
11	311	2.397895	5.739793	13.76342	5.749902
12	236	2.484907	5.463832	13.57711	6.174761
13	236	2.564949	5.463832	14.01445	6.578965
14	230	2.639057	5.438079	14.3514	6.964624
15	177	2.70805	5.17615	14.01727	7.333536
16	117	2.772589	4.762174	13.20355	7.687248
17	118	2.833213	4.770685	13.51637	8.027098
18	59	2.890372	4.077537	11.7856	8.354249
19	57	2.944439	4.043051	11.90452	8.669721
20	52	2.995732	3.951244	11.83687	8.974412
21	51	3.044522	3.931826	11.97053	9.269117
22	51	3.091042	3.931826	12.15344	9.554543
23	110	3.135494	4.70048	14.73833	9.831324
24	101	3.178054	4.615121	14.6671	10.10003
25	60	3.218876	4.094345	13.17919	10.36116
31	30	3.433987	3.401197	11.67967	11.79227
Total	57640	61.43759	147.1721	311.4875	163.0732

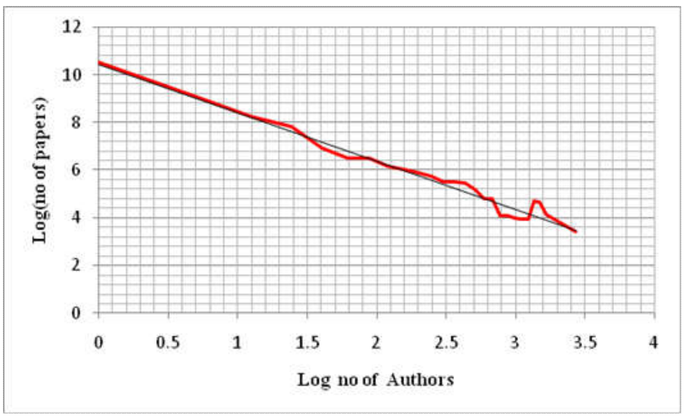

Figure 1. Fitted linear line for authorship distribution in the field of Neuroscience

Table-3 Fitted Lotka's Distribution with n=2.02702 and with n=2.05681

No of papers	Number of authors observed	Expected with LSn=2.02702	Standardized value(n=2.02702)	Expected with ML=2.05681	Standardized value(n=2.05681)
1	36803	36803	36336.66	36803	36837.79
2	9074	9030.034	8915.611	8845.486	8853.848
3	3761	3969.62	3919.319	3841.806	3845.438
4	2440	2215.622	2187.547	2125.985	2127.995
5	1007	1409.474	1391.614	1343.491	1344.761
6	664	973.9913	961.6495	923.3661	924.239
7	652	712.6111	703.5814	672.4765	673.1122
8	474	543.6279	536.7394	510.9739	511.4569
9	414	428.1684	422.7429	401.04	401.4191
10	355	345.8304	341.4483	322.9038	323.2091
11	311	285.0752	281.4629	265.4216	265.6726
12	236	238.9798	235.9516	221.9282	222.138
13	236	203.1879	200.6132	188.2406	188.4186
14	230	174.8472	172.6317	161.6276	161.7804
15	177	152.0277	150.1013	140.2449	140.3774
16	117	133.3853	131.6951	122.811	122.9271
17	118	117.9611	116.4663	108.4136	108.516
18	59	105.056	103.7248	96.38871	96.47983
19	57	94.15082	92.9578	86.24423	86.32576
20	52	84.85343	83.77822	77.60893	77.6823
21	51	76.86317	75.8892	70.19875	70.26511
22	51	69.94644	69.06012	63.79326	63.85357
23	110	63.91954	63.1096	58.21941	58.27445
24	101	58.63642	57.89341	53.33975	53.39017
25	60	53.97975	53.29575	49.04404	49.0904
31	30	34.90304	34.46077	31.50907	31.53886
Total	57640	58379.75	57640	57585.56	57640

Table-4 Kolmogorov Smirnov test for (n=2.0702)

No of Papers	Observed Frequency of Authors	Observed Cumulative Frequency of Authors	Theoretical Frequency of Authors	Theoretical Frequency	Theoretical cumulative Frequency	Difference
1	36803	0.638498	0.638498	0.6074	0.6074	0.031098
2	9074	0.157425	0.795923	0.15185	0.75925	0.036673
3	3761	0.06525	0.861173	0.067489	0.826739	0.034434
4	2440	0.042332	0.903505	0.037963	0.864701	0.038804
5	1007	0.017471	0.920975	0.024296	0.888997	0.031978
6	664	0.01152	0.932495	0.016872	0.90587	0.026626
7	652	0.011312	0.943807	0.012396	0.918266	0.025541
8	474	0.008223	0.95203	0.009491	0.927756	0.024274
9	414	0.007183	0.959213	0.007499	0.935255	0.023958
10	355	0.006159	0.965372	0.006074	0.941329	0.024043
11	311	0.005396	0.970767	0.00502	0.946349	0.024419
12	236	0.004094	0.974862	0.004218	0.950567	0.024295
13	236	0.004094	0.978956	0.003594	0.954161	0.024795
14	230	0.00399	0.982946	0.003099	0.95726	0.025686
15	177	0.003071	0.986017	0.0027	0.959959	0.026058
16	117	0.00203	0.988047	0.002373	0.962332	0.025715
17	118	0.002047	0.990094	0.002102	0.964434	0.02566
18	59	0.001024	0.991118	0.001875	0.966309	0.024809
19	57	0.000989	0.992107	0.001683	0.967991	0.024116
20	52	0.000902	0.993009	0.001519	0.96951	0.023499
21	51	0.000885	0.993894	0.001377	0.970887	0.023007
22	51	0.000885	0.994778	0.001255	0.972142	0.022637
23	110	0.001908	0.996687	0.001148	0.97329	0.023397
24	101	0.001752	0.998439	0.001055	0.974345	0.024094
25	60	0.001041	0.99948	0.000972	0.975316	0.024164
31	30	0.00052	1	0.000632	0.975948	0.024052
Total	57640	1				

of the total publications, 9074 authors contributed only two papers (16%) and three papers are by 6.5 % of the authors and so on. The observed values are almost equal to the theoretical value for the give data set, from the above discussion it can be concluded that single-authored papers are more dominant than the multi-authored papers in neuroscience research. The estimated value of n for the give data set is calculated by using the equation -1 the n value in the field of Neuroscience is -2.05681 for all author data.

The table -2 indicates the calculation of exponent 'n' for overall author productivity in the field of Neuroscience.

$$\frac{26\sum 311.4875 - 61.4375 * 147.1721}{26 * 163.0732 - (61.43759)^2}$$

n=-2.05681

The table-3 shows the standardized fitted distribution using estimated value LSn= 2.02702 and ML=2.05681for the given data set in the field of Neuroscience. It is observed that the observed values and theoretical values are almost similar but there is slight variation in the standardized value. This shows that given Lotka's law of author productivity is almost applicable to Neuroscience literature.

Goodness- of- fit test: The observed and known data were tested for goodness-of-fit using the one-sample Kolmogorov-Smirnov(K-S) test. The K-S goodness-of-fit test was used to compare the functions describing the observed and theoretical distributions of publications. Early studies have used either the Chi-square or the K-S goodness-of-fit test when testing data conformity to Lotka's law. However, recent studies (Budd & Seavey, 1990; Cline, 1981; Colie, 1977; Pao, 1985) have employed the K-S test because it is seen as the most powerful test available (Black, 2003; Nicholls, 1986; Pao, 1985). Black (2003) notes the K-S test is more powerful than the 2-test and is an appropriate test for ranked data. Since one of the goals of this study is to prove that the observed distribution is not different from the distribution predicted by Lotka,

The Critical value is calculated by using the formula suggested by Black

The critical value is $D=\frac{1.3\ \ 6}{\sqrt{\sum N+1}}=\frac{1.3\ \ 6}{\sqrt{3\ \ 26}}=0.075439$

The value of n and c can be calculated, to be n= 2.02702 and c= .630327 respectively. The Kolmogorov Smirnov test (KS) one sample goodness of fit test was conducted at the 1% level of significance, the Dmax is 0.038804 and the resulted critical value is 0.075439. Since the critical value is greater than the D max value (0.038804). We must fail to reject null hypothesis i.e the observed Authorship data Distribution is the same as the Theoretical Authorship Distribution. It can be concluded that the given data set well fitted in the Lotka's distribution of author productivity in the field of Neurosciences. (table-4).

Conclusion

Lotka's law is essentially to predict the author productivity in the given field. Hence, it cannot be used for the productivity of the individual authors. The present study has provided an overview of author productivity in the field of Neuroscience literature published from 2000 to 2018. In the present study value of n and c are on a little higher side implying that the Neuroscience field has relatively increased. It was also observed that if Pao's methodology is applied correctly on proper sample size in terms of the unique number of author and coverage, it would give better results than the methodology suggested by Nicholls. Finally, it can be concluded that Lotka's law of author productivity is well applicable to Neuroscience literature.

REFERENCES

Allan Piechard 1969. A Statistical Bibliography an interim bibliography., New Western Polytechnic School of Librarianship, London p.14.

Askew C A. 2008. An examination of Lotka'a law in the field of library and information studies. FIU Electronic Theses and Dissertations. p.182 available on http//digital common FIU edu/etd/182 accessed on 28.11.2018.

Burrell QL. 2004. Fitting Lotka's law: some cautionary observations on a recent paper by Newby et. Al. (2003), Journal of the American Society for Information Science and Technology, 55(13), 1209-1211.

Burrell Q L. 1994. The Kolmogorov- Smirnov test and rank-frequency distribution, Journal of the American Society for Information Science, 45(1) 50.

Coile R C. 1977. Lotka's frequency distribution of scientific productivity, Journal of the American Society for Information Science, 28(6) 366-370.

Lotka A J. 1926. The frequency distribution of scientific productivity, Journal of the

Washington Academy of Sciences, 16(2) 317-323. *Management*, 21(4)(1985) 305-320.

Nicholls P T. 1989. Bibliometric modeling processes and the empirical validity of Lotka's Law. Journal of American Society for Information Science, 40(6) 379-385.

Newby G B. 2003. Greenberg J and Jones P, Open source Software development and Lotka's law: Bibliometric patterns in programming, Journal of the American Society Information Science and Technology, 54(2) 169-178.

Pao. M. L. 1985. Lotka's law: A testing procedure, information processing, and management, 21(4). 305-320.

Peter M. 2008. Personal name heading in COBIB: testing Lotka's law, Scientometrics, 75(1) 175-188.

****