2019

Application of Lotka's Law in Bell's palsy (facial paralysis) research output during 2004 -2018.

K.T Naheem Mr
Pondicherry University, naheemkt@gmail.com

P Sivraman Dr.
Annamalai University, psraman.p@gmail.com

G Saravanan Dr.
French Institute of Pondicherry, saravanan.g@ifpindia.org

Application of Lotka's Law in Bell's palsy *(facial paralysis)* research output during 2004 - 2018.

Naheem K.T.,[1]Research Scholar, Dept. of Library & Information Science, Annamalai University, Annamalai Nagar – 608002.; [1]Professional Assistant, Central Library, Pondicherry University, Kalapet, Puducherry-605014. Email: naheemkt@gmail.com;

Dr. P. Sivaraman[2]., Associate Professor, Dept. of Library & Information Science, Annamalai University, Annamalai Nagar – 608002. Email: psraman.p@gmail.com

Dr. G. Saravanan[3]., Librarian, Library and Publication Division, French Institute of Pondicherry. Email: saravanan.g@ifpindia.org

ABSTARCT

This paper examines the validity of Lotka's law of scientific publication productivity of the articles published on Bell's palsy disease during 2004-2018. Data for this analysis was retrieved from WOS data base of Clarivate analytics. In this study, the straight and complete count of authorship was used. A total of 4039 articles along with 3384 and 14517 authors were identified by using straight and complete count method of authorship respectively. K-S goodness of- fit statistical test were employed to verify the applicability of Lotka's law. The results showed that, Lotka's law fits with the data of straight count of authors. While this law doesn't find fits to complete count authorships. Hence, it is concluded that Lotka's law partially fits with Bell's palsy literature.

Keywords: Lotka's law, Bell's palsy, Authorship distribution, Kolmogorov-Smirnov test, Bibliometrics**.**

1. INTRODUCTION

Lotka's Law is treated as one of the prominent laws of bibliometrcis. In 1926 Alfred J. Lotka[1] proposed a law which is often called "inverse square law" indicating that there is an inverse relation between the number of publications and the number of authors producing these publications. Lotka's Law describes the frequency of publication by authors in a given subject field. It stated that:

> "... the number of authors making n contributions is about $1/n^2$ of those making one; and the proportion of all contributors, that make a single contribution, is about 60 percent".

This means that out of all the authors in a given subject field, about 60% publish only one article, 15% publish two articles, 7% publish three articles, and so on. According to Lotka's Law, only 6% authors in a subject field produce more than ten articles. The generalized form of Lotka's Law can be expressed as "$x^n y = c$", where y is the number of authors with x articles, the

exponent *n* and constant *c* are parameters to be estimated from a given set of author productivity data.

The validity of Lotka's law has been studied by a number of researchers who have applied the model to data sets in many subject areas. Most notable are the contribution of Pao[2] and Nicholls[3], who found that the Lotka model fitted the majority of the data sets studied [5]. The present study is an attempt to apply Lotka's Law and test its conformity on the authorship distribution in the literature on Bell's palsy disease using WOS data during 2004-2018.

2. LITERATURE REVIEW

Zabed and Rahman[4] examined the validity of Lotka's Law to authorship distribution in the field of Nutrition Research in Bangladesh .The results of this study showed that author productivity distribution predicted in Lotka's generalized inverse square law is not applicable to Nutrition Research of Bangladesh. While, using least-squares excluding high productive authors Lotka's Law was found to be applicable to Nutrition Research of Bangladesh. Sudheer[5] applied lotka low to physics literature and reported that the distribution author productivity doesn't fit Lotka's low to physics literature. Askew[6] examined the conformity of Lotka's law in the Field of Library and Information Studies. Results of this study confirmed the validity of Lotka's law for predicting author productivity in the field of library and information studies. Kumar and Senthilkumar[7] applied and tested the conformity of Lotka's law in Astronomy & Astrophysics literature. Reported that, Lotka's law doesn't fits to the literature of Astronomy & Astrophysics. Suresh Kumar[8] verified the authorship frequency distribution on 2106 publications references of the Journal of Documentation published during 2003-2015 using K-S test and conculded that Lotka's law fit to the data set of LIS literature. Sevukan and Sharma[9] studied the biotechnology research in central universities of India and found that the Lotka's law fits to the data of Biotechnology literature. Naqvi and Fatima[10] studied applicability of Lotka's law in international business literature and found that Lotka's distribution is applicable to international business literature. Kumar[11] analysed the distribution of author productivity in the Human-Computer Interaction (HCI) research output and reported the conformity of Lotka's law to HCI literature.

3. OBJECTIVES

1. To examine the validity of Lotka's law using straight and complete count of authors; and
2. To apply Kolmogorov- Smirnov (K-S) goodness-of-fit test for the conformity of Lotka's law.

4. METHODOLOGY

The data for the study were retrieved from Clarivate analytics WOS Science Citation Index Expanded database. The terms, 'Bell's palsy" and "facial paralysis" were used for search in the topic filed. Document type 'articles' published during 2004 to 2018 were identified and used for

further analysis. The straight and complete counts of authors were considered for the present analysis. Kolmogorov- Smirnov (K-S) goodness-of-fit test were employed for the conformity of Lotka's law.

The final data consists 4039 articles along with 3384 authors (straight count) and 14517 authors (complete count). The analysis were carried out using the following steps: calculation of n, c, K-S test and calculation critical value.

a.Estimation of the exponent n

The value of *n* is calculated by the least square-method using the following formula:

$$n = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - \left(\sum X\right)^2} \quad \ldots\ldots (1)$$

N = number of pairs of data
X = logarithm of x, i. e. number of articles
Y = logarithm of y, i. e. number of authors

b.Calculation of the constant c

The value of *c* is estimated using the following formula:

$$\sum \frac{1}{x^n} \quad \ldots\ldots (2)$$

c. Kolmogorov-Smirnov (K-S) test

K-S test, a goodness-of-fit statistical test is applied to state that the observed author productivity distribution is not significantly different from a theoretical distribution. The maximum deviation between the cumulative proportions of the observed and theoretical frequency is determined by the following formula:

$$D = \max\left|F_0(x) - S_n(x)\right| \quad \ldots\ldots (3)$$

$F_0(x)$ = theoretical cumulative frequency
$S_n(x)$ = observed cumulative frequency

d. Calculation of critical value.

The critical value at the 0.01 level of significance is calculated using the following equation:

$= 1.63 / [\sum yx + (\sum yx/10)^{1/2}]^{1/2}$ ..... (4)

$\sum yx$ = the total population under study

5. RESULTS AND DISCUSSIONS

5.1 Analysis of data based on Straight count method

(a) Estimation of the parameters *n* and *c*

The first step in the application of Lotka's law is to determine the value of *n*, which is to be determined by Linear Least Square (LLS) method by the mathematical formula (1). To calculate the parameter 'n' data given in column 1 to 7 in the Table 1 is used.

By substituting the values in equation (4), the value of n is calculated as :

$$n = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - \left(\sum X\right)^2} = \frac{13(4.6622) - (10.5786)(11.8498)}{13(112.0977) - (10.5876)^2} = -2.016$$

The second step in the application of Lotka's law is to determine the value of 'c'. For which data given in column 9 in Table 1 is used. By replacing the values in the equation (2), value of is obtained as:

$$\sum \frac{1}{x^n} = \sum \frac{1}{1.5435} = 0.64$$

Table -1: Calculation of n and c–Straight count method

Sl.no	x	yx	X- Log x	Y-Log y	XY	X2	xn	1/xn
Col.1	*Col.2*	*Col.3*	*Col.4*	*Col.5*	*Col.6*	*Col.7*	*Col.8*	*Col.9*
1	1	3011	0.00000	3.47871	0.00000	0.00000	1.00000	1.00000
2	2	252	0.30103	2.40140	0.72289	0.09062	4.04461	0.24724
3	3	67	0.47712	1.82607	0.87126	0.22764	9.15960	0.10918
4	4	29	0.60206	1.46240	0.88045	0.36248	16.35886	0.06113
5	5	10	0.69897	1.00000	0.69897	0.48856	25.65214	0.03898
6	6	4	0.77815	0.60206	0.46849	0.60552	37.04699	0.02699
7	7	3	0.84510	0.47712	0.40321	0.71419	50.54959	0.01978
8	8	2	0.90309	0.30103	0.27186	0.81557	66.16517	0.01511
9	9	1	0.95424	0.00000	0.00000	0.91058	83.89825	0.01192
10	14	2	1.14613	0.30103	0.34502	1.31361	204.45330	0.00489
11	16	1	1.20412	0.00000	0.00000	1.44990	267.61219	0.00374

12	17	1	1.23045	0.00000	0.00000	1.51400	302.40226	0.00331
13	28	1	1.44716	0.00000	0.00000	2.09427	826.93351	0.00121
	Total	**3384**	**10.58762**	**11.84983**	**4.66216**	**112.09765**	**1895.27644**	**1.54348**

n=2.016, c=0.64

(b) Application of K-S Statistical test

The K-S test is performed to examine the conformity of the observed author distribution versus Lotka's distribution in straight count method. For applying K-S test, convert the observed and expected number of authors into fractional values, and take the difference between cumulative fractional values of observed and expected number of authors. K-S test for straight count method is carried out with the help of the following Table 2.

Table 2. K-S test - Straight count method *(n* = 2.016)

		Observed		**Expected**		
x	**yx**	**% of authors**	**Cumulative % of authors**	**% of authors**	**Cumulative % of authors**	**Difference**
col.1	***col.2***	***col.3***	***col.4***	***col.5***	***col.6***	***(col.4-col.6)***
1	3011	0.88978	0.88980	0.64000	0.64000	0.24980
2	252	0.07447	0.96427	0.15824	0.79824	0.16603
3	67	0.01980	0.98407	0.06987	0.86811	0.11596
4	29	0.00857	0.99264	0.03912	0.90723	0.08541
5	10	0.00296	0.99559	0.02495	0.93218	0.06341
6	4	0.00118	0.99677	0.01728	0.94945	0.04732
7	3	0.00089	0.99766	0.01266	0.96212	0.03555
8	2	0.00059	0.99825	0.00967	0.97179	0.02646
9	1	0.00030	0.99855	0.00763	0.97942	0.01913
14	2	0.00059	0.99914	0.00313	0.98255	0.01659
16	1	0.00030	0.99943	0.00239	0.98494	0.01450
17	1	0.00030	0.99973	0.00212	0.98705	0.01267
28	1	0.00030	1.00002	0.00077	0.98783	0.01220
Total	3384				**Dmax**	**0.24980**

cv-2.277

$$D = \max|F_0(x) - S_n(x)| = 0.24980$$

The critical value at the 0.01 level of significance:

$=1.63 /[\sum yx + (\sum yx/10)^{1/2}]^{1/2}$

= 1.63/ 58.67003

=0.27782

It is evident from Table 2, that the maximum deviation is identified as (0.249) which is lower than the calculated critical value (0.27782) at the 0.01 level of significance. Therefore, the test confirming that the distribution of the data using straight count author fit's Lotka's law.

6.2 Analysis of data based on Whole author count method

(a) Estimation of the parameters *n* and *c*

The value of '*n'*, related to whole count method is calculated with the mathematical formula (1). To calculate the parameter 'n' data given in column 1 to 7 in the Table 2 is used.

By substituting the values in equation (4), the value of n is calculated as :

$$n = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - \left(\sum X\right)^2} = \frac{27(17.6467) - (29.0022)(26.8753)}{27(35.5291) - (29.0022)^2} = -2.56$$

The value of the parameter '*c*' is determined by data given in column 9 in Table 3. By replacing the values in the equation (2), value of '*c*' is obtained as:

$$\sum \frac{1}{x^n} = \sum \frac{1}{1.3147} = 0.76$$

Table 3. Calculation of n and c—whole authors

Sl.no	x	y	X (Log x)	Y(Log y)	XY	X^2	x^n	1/xn
1	1	11954	0.0000	4.0775	0.0000	0.0000	1.0000	1.0000
2	2	1530	0.3010	3.1847	0.9587	0.0906	5.8971	0.1696
3	3	473	0.4771	2.6749	1.2762	0.2276	16.6506	0.0601
4	4	218	0.6021	2.3385	1.4079	0.3625	34.7755	0.0288
5	5	128	0.6990	2.1072	1.4729	0.4886	61.5692	0.0162
6	6	59	0.7782	1.7709	1.3780	0.6055	98.1900	0.0102
7	7	44	0.8451	1.6435	1.3889	0.7142	145.6971	0.0069
8	8	24	0.9031	1.3802	1.2465	0.8156	205.0739	0.0049
9	9	22	0.9542	1.3424	1.2810	0.9106	277.2431	0.0036
10	10	21	1.0000	1.3222	1.3222	1.0000	363.0781	0.0028
11	11	8	1.0414	0.9031	0.9405	1.0845	463.4099	0.0022
12	12	5	1.0792	0.6990	0.7543	1.1646	579.0339	0.0017

13	15	5	1.1761	0.6990	0.8221	1.3832	1025.1647	0.0010
14	16	5	1.2041	0.6990	0.8416	1.4499	1209.3365	0.0008
15	13	3	1.1139	0.4771	0.5315	1.2409	710.7142	0.0014
16	17	3	1.2304	0.4771	0.5871	1.5140	1412.3725	0.0007
17	19	3	1.2788	0.4771	0.6101	1.6352	1877.6266	0.0005
18	20	2	1.3010	0.3010	0.3916	1.6927	2141.0992	0.0005
19	28	2	1.4472	0.3010	0.4356	2.0943	5066.6926	0.0002
20	14	1	1.1461	0.0000	0.0000	1.3136	859.1871	0.0012
21	18	1	1.2553	0.0000	0.0000	1.5757	1634.9242	0.0006
22	24	1	1.3802	0.0000	0.0000	1.9050	3414.6074	0.0003
23	25	1	1.3979	0.0000	0.0000	1.9542	3790.7612	0.0003
24	27	1	1.4314	0.0000	0.0000	2.0488	4616.2704	0.0002
25	30	1	1.4771	0.0000	0.0000	2.1819	6045.4748	0.0002
26	46	1	1.6628	0.0000	0.0000	2.7648	18057.6009	0.0001
27	66	1	1.8195	0.0000	0.0000	3.3107	45502.2084	0.0000
Total		**14517**	**29.0022**	**26.8753**	**17.6467**	**35.5291**	**99615.6587**	**1.3147**

n-2.56, c-0.76

(b) Application of K-S Statistical test

KS test for whole authors is carried out with the help of the following Table.4.

Table 4. K-S test–whole authors *(n* = 2.56)

		Observed		**Expected**		
x	**yx**	**% of authors**	**Cumulative % of authors**	**% of authors**	**Cumulative % of authors**	**Difference**
col.1	***col.2***	***col.3***	***col.4***	***col.5***	***col.6***	***(col.4-col.6)***
1	11954	0.82345	0.82340	0.76000	0.76000	**0.06340**
2	1530	0.10539	0.92879	0.12888	0.88888	0.03992
3	473	0.03258	0.96138	0.04564	0.93452	0.02685
4	218	0.01502	0.97639	0.02185	0.95638	0.02002
5	128	0.00882	0.98521	0.01234	0.96872	0.01649
6	59	0.00406	0.98927	0.00774	0.97646	0.01281
7	44	0.00303	0.99231	0.00522	0.98168	0.01063
8	24	0.00165	0.99396	0.00371	0.98538	0.00858

9	22	0.00152	0.99547	0.00274	0.98812	0.00735
10	21	0.00145	0.99692	0.00209	0.99022	0.00670
11	8	0.00055	0.99747	0.00164	0.99186	0.00562
12	5	0.00034	0.99782	0.00131	0.99317	0.00465
15	5	0.00034	0.99816	0.00074	0.99391	0.00425
16	5	0.00034	0.99851	0.00063	0.99454	0.00397
13	3	0.00021	0.99871	0.00107	0.99561	0.00310
17	3	0.00021	0.99892	0.00054	0.99615	0.00277
19	3	0.00021	0.99913	0.00040	0.99655	0.00257
20	2	0.00014	0.99926	0.00035	0.99691	0.00236
28	2	0.00014	0.99940	0.00015	0.99706	0.00234
14	1	0.00007	0.99947	0.00088	0.99794	0.00153
18	1	0.00007	0.99954	0.00046	0.99841	0.00113
24	1	0.00007	0.99961	0.00022	0.99863	0.00098
25	1	0.00007	0.99968	0.00020	0.99883	0.00085
27	1	0.00007	0.99974	0.00016	0.99899	0.00075
30	1	0.00007	0.99981	0.00013	0.99912	0.00070
46	1	0.00007	0.99988	0.00004	0.99916	0.00072
66	1	0.00007	0.99995	0.00002	0.99918	0.00077
	14517	1.00000		0.99918	**D *max***	**0.06340**

D max-0.06340, c.v. 0.01347

$$D = \max|F_0(x) - S_n(x)| = 0.06340$$

The critical value at the 0.01 level of significance: $=1.63 /[\sum yx + (\sum yx/10)^{1/2}]^{1/2}$

c.v = 1.63/ 120.9854

c.v =0.01347

It is observed from Table 4, that the maximum deviation is identified as 0.063. The critical value D in K-S test at 0.01 level of significance is 0.01347. While comparing these values, it is found that the actual value of *D (0.063)* does not fall within the critical value of *D (0.01347).* Therefore, the test confirming that the distribution of the data using whole author count doesn't fit Lotka's law.

CONCLUSION

Lotka's Law of scientific productivity is regarded as one of the classical laws of bibliometrics. An important area of bibliometrics is author productivity, which can be measured by a number of publications of a particular author. The present analysis revealed that the Lotka's law in its

modified form fit the author productivity distribution pattern prepared for the straight count by based on the calculated values of 'n' (2.016) and 'c' (0.64). At the same, it is found that, the author productivity distribution pattern prepared for the complete count doesn't fits. This study is a maiden attempt to analyse the authorship distribution in the literature on Bell's palsy diseases. Hence the results may be used for future in-depth studies on similar kind of subject areas.

REFERENCES

1. Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington academy of sciences*, *16*(12), 317-323.
2. Pao, M. L. (1986). An empirical examination of Lotka's law. *Journal of the American Society for Information Science*, *37*(1), 26-33.
3. Nicholls, P. T. (1989). Bibliometric modeling processes and the empirical validity of Lotka's law. *Journal of the American Society for Information Science*, *40*(6), 379-385.
4. AhmedI, S. Z., & Rahman, M. A. (2009). Lotka's law and authorship distribution innutrition research in Bangladesh. *Annals of Library and Information Studies*, *56*, 95-102.
5. Pillai Sudhier, K. G. (2013). Lotka's Law and Pattern of Author Productivity in the Area of Physics Research. *DESIDOC Journal of Library & Information Technology*, *33*(6), 457-464.
6. Askew, C. A. (2008) An Examination of Lotka's law in the Field of Library and Information Studies. *FIU ElectronicTheses and Dissertations*,182.
7. Kumar, S., & Senthilkumar, R. (2019). Applicability of Lotka's Law in Astronomy & Astrophysics Research of India. *Library Philosophy and Practice*, 1-13..http://digitalcommons.unl.edu/libphilprac/2129
8. Suresh Kumar, P.K. (2018). Author productivity and the application of Lotka's Law in LIS publications. *Annals of Library and Information Studies (ALIS)*, *64*(4), 234-241.
9. Sevukan, R., & Sharma, J. (2008). Bibliometric analysis of research output of biotechnology faculties in some Indian central universities. *DESIDOC Journal of Library & Information Technology*, *28*(6), 11-20.
10. Naqvi, S. H., & Fatima, N. (2018). Authorship patterns in international business literature: applicability of Lotka's Law. *Annals of Library and Information Studies (ALIS)*, *64*(4), 253-259.
11. Kumar, S. (2015). Author productivity in the field Human Computer Interaction (HCI) research. *Annals of Library and Information Studies (ALIS)*, *61*(4), 273-285.