Lotka's Law & Productivity trends: An analysis

Dr. Jagdish Narharrao Kulkarni*

* **Librarian, (CAS),** Knowledge Resource Center, Swami RamanandTeerth Marathwada University, Nanded, Maharashtra India.

QR Code

Abstract: - *In the area of scholarly communication one is always curious to know the publication productivity trend. Alfred J. Lotka proposed his inverse square law correlating contributions of scientific papers to their number of contributions. The study covered 21 proceedings of CALIBER, NACLIN & IASLIC from 2003 to 2009 in which 1706 papers were published. Attempt has been made to examine the Lotka's law to the set of data of conference proceedings.*

Keywords: Scientometrics, Bibliometrics, Productivity trend, Lotka's Law.

1. Introduction:

In 1926, Alfred J. Lotka proposed his inverse square law correlating contributions of scientific papers to their number of contributions. His law provided fundamental theoretical base for bibliometric studies involving authorships.

In the case examined it is found that the number of persons making two contributions is about one-fourth of those making one contribution, the number making 'n' contributions is about $1/n^2$ of those making one.

In other words, for every 100 authors contributing one article, 25 will contribute two articles, about 11 will contribute 3 articles and 6 will contribute 4 articles and so on. Though the law is based on the study of chemistry and physics literature later it has generated much interest and attracted the attention of researchers and it has been applied and tested in many other fields (Sangam, 2008).

2. Productivity trends and application of Lotka's Law

Productivity trends for well-established and recognized subjects and discipline of universal nature like physics, chemistry and Mathematics etc. follow distribution pattern which confirms Lotka's Law, if applied to such distributions in it's original form with exponent value of two (Vaishnav&Aghav, 1994). Several studies for

scientific disciplines have been reported in the literature of Library and Information Science. It was thought that it would be appropriate and useful to investigate the applicability of Lotka's Law to the present set of data. The data from present study on author productivity is presented in table 1.

Table 1: Productivity of authors based on Lotka's law

	Overall		CALIBER		NACLIN		IASLIC	
No. of Papers	**Observed no. of authors**	**Expected no. of authors**	**Observed no. of authors**	**Expected no. of authors**	**Observed no. of authors**	**Expected no. of authors**	**Observed no. of authors**	**Expected no. of authors**
1	722	722	474	474	161	161	308	308
2	168	180.5	105	118.5	24	40.25	51	77
3	69	80.222	29	52.666	8	17.889	17	34.222
4	32	45.125	6	29.625	2	10.0625	18	19.25
5	20	28.88	4	18.96	2	6.44	7	12.32
6	10	20.056	2	13.167	0	4.472	4	8.555
7	10	14.735			2	3.285	2	6.285
8	3	11.281					1	4.812
9	3	8.913						
10	2	7.22						
11	0	5.966						
12	1	5.014						
	1040	722	**620**		**199**		**408**	

It can be noted from the table 1 that the productivity distribution does not fit Lotka's inverse square law applied to it, in overall as well as individual data sets of CALIBER, NACLIN & IASLIC.

3. Determination of exponent value (α)

Since the productivity distribution has not fitted Lotka's inverse square law applied to it; hence a different procedure was adopted.

The value of productivity constant was determined by using following formula for the data when full authorship was given only to first authors:

$$n^{\alpha} = \frac{\text{Proportion of authors contributing n papers}}{\text{Proportion of authors contributing only 1 paper}}$$

Where n= 1, 2, 3, 4, 5. . .12 andα is the productivity constant or characteristic exponent for overall data. The value of n was n= 1, 2, 3, 4, 5, 6 for CALIBER, n= 1, 2, 3, 4, 5, 6, 7; for NACLIN and n= 1, 2, 3, 4, 5, 6, 7, 8 for IASLIC set of data.

Average value of α was found to be 2.1494 for overall data. This value can be approximated to be as 2.15. Calculated average values of α and its approximated values (shown in circular bracket) were 2.77762 (2.77) for CALIBER; 2.4196 (2.42) for NACLIN and 2.48561 (2.48) for IASLIC.

4. Determination of estimated proportion of authors

Having found the value of α, Lotka's fraction $1/ n^{\alpha}$ was summed up for all values of N= α applying Euler-Maclauring formula of summation. Then the sum was used as a divisor for $1/ n^{\alpha}$ to determine the proportion of the total number of authors who should be expected to produce n papers. Following formula was used to find the proportions,

First the value of S was calculated by using formula,

$$S = \mathbf{\Sigma 1/n^{\alpha}} = \sum_{n=1}^{12} 1/ n^{2.15} = \mathbf{1.47615} \text{ ---------- (1)}$$

Where n= 1, 2, 3, 4, 5. . .12 for overall set of data;

$$S = \mathbf{\Sigma 1/n^{\alpha}} = \sum_{n=1}^{6} 1/ n^{2.77} = \mathbf{1.234356} \text{ ---------- (2)}$$

Where n= 1, 2, 3, 4, 5, 6 for CALIBER set of data;

$$S = \mathbf{\Sigma 1/n^{\alpha}} = \sum_{n=1}^{7} 1/ n^{2.42} = \mathbf{1.334262} \text{ ---------- (3)}$$

Where n= 1, 2, 3, 4, 5, 6, 7 for NACLIN set of data and

$$S = \mathbf{\Sigma 1/n^{\alpha}} = \sum_{n=1}^{8} 1/ n^{2.48} = \mathbf{1.320954} \text{---------- (4)}$$

Where n = 1, 2, 3, 4, 5, 6, 7, 8 for IASLIC set of data.

For present study where S is the sum of Lotka's modified ratio for the values of α which is equal to 1.47615for overall data calculated by equation 1. Further the value of S calculated for CALIBER (equation 2), NACLIN (equation3) and IASLIC (equation 4), which was found to be 1.234356 for CALIBER, 1.334262 for NACLIN and 1.320954for IASLIC set of data.

The expected number of authors (An) was calculated for present set of overall data and for CALIBER, NACLIN and IASLIC set of data by using the formula,

$$An = \frac{1/n^{\alpha}}{S} * T$$

Where ,α is the productivity constant or characteristic exponent

T is total number of authors in the sample and

An is the total number of expected authors producing n papers.

Where n = 1, 2, 3, 4, 5. . .12 for overall set of data;

n = 1, 2, 3, 4, 5, 6 for CALIBER;

n = 1, 2, 3, 4, 5, 6, 7 for NACLIN and

n = 1, 2, 3, 4, 5, 6, 7, 8 for IASLIC set of data.

The values of An for overall as well as CALIBER, NACLIN and IASLIC proceedings are shown in table 2 for overall data, table 3 for CALIBER data, table 4 for NACLIN and table 5 for IASLIC set of data.

5. Application of statistical tests

After the values of α, Sand proportion of authors (An) were determined; the observed and estimated values of the proportions were statistically tested by applying K-S test to present set of data. The data for K-S test with the value of α = 2.15 for overall data are specified in table 2.

Table 2 Productivity Trend: Proportion of author (Overall)

No. of contribution	No. of Authors Observed	Observed	Sn(x)	No. of Authors Expected (An)	Expected	Fo(x)	|Fo(x) - Sn(x)|	
1	722	0.6942	0.6942	704.54	0.6774	0.6774	0.016800	
2	168	0.1615	0.8557	158.74	0.1526	0.8300	0.025703	
3	69	0.0663	0.9221	66.39	0.0638	0.8939	**0.028214**	**Maximum**
4	32	0.0308	0.9529	35.77	0.0344	0.9283	0.024593	
5	20	0.0192	0.9721	22.14	0.0213	0.9495	0.022538	
6	10	0.0096	0.9817	14.96	0.0144	0.9639	0.017771	
7	10	0.0096	0.9913	10.74	0.0103	0.9743	0.017061	
8	3	0.0029	0.9942	8.06	0.0077	0.9820	0.012197	
9	3	0.0029	0.9971	6.26	0.0060	0.9880	0.009066	
10	2	0.0019	0.9990	4.99	0.0048	0.9928	0.006193	
11	0	0.0000	0.9990	4.06	0.0039	0.9967	0.002286	
12	1	0.0010	1.0000	3.37	0.0032	1.0000	0.000007	
	1040	1.0000		1040.00	1.0000			

Table 2 depicts Productivity Trend with proportion of authors for overall proceedings where only first authors were considered with exponent value of $\alpha = 2.15$

Dmax = 0.028214

Dmax = |Fo(x) -Sn(x)| = 0.028214

At 0.01 level of significance, K. S. Static = 1.63/sqrt (1040) = 0.050544

Dmax = 0.028214 <0.050544, therefore data fits in to generalized form of Lotka's law with exponent value of $\alpha = 2.15$.

Further attempt has been made to test applicability of Lotka's law to the individual data set of CALIBER, NACLIN and IASLIC conference proceedings and shown in table 3, 4 and 5respectively.

Table 3 Productivity Trend: Proportion of author (CALIBER)

No. of contribution	**No. of Authors Observed**	**Observed**	**Sn(x)**	**No. of Authors Expected (An)**	**Expected**	**Fo(x)**	**|Fo(x) - Sn(x)|**	
1	474	0.7645	0.7645	502.286	0.8101	0.8101	**0.045600**	**Maximum**
2	105	0.1694	0.9339	73.6374	0.1188	0.9289	0.004985	
3	29	0.0468	0.9806	23.9511	0.0386	0.9675	0.013128	
4	6	0.0097	0.9903	10.7956	0.0174	0.9849	0.005393	
5	4	0.0065	0.9968	5.81841	0.0094	0.9943	0.002460	
6	2	0.0032	1.0000	3.51134	0.0057	1.0000	0.000023	
	620	1.0000		**620**	1.0000			

Table 3 depicts Productivity Trend with proportion of authors for CALIBER where only first authors were considered with exponent value of α= 2.77.

Dmax =0.0456

Dmax = |Fo(x) -Sn(x)| = 0.0456

At 0.01 level of significance, K. S. Static = 1.63/sqrt (620) = 0.065462

Dmax = 0.0456<0.065462, therefore data fits in to generalized form of Lotka's law with exponent value of α= 2.77.

Table 4 Productivity Trend: Proportion of author (NACLIN)

No. of contribution	No. of Authors Observed	Observed	Sn(x)	No. of Authors Expected (An)	Expected	Fo(x)	|Fo(x) - Sn(x)|	
1	161	0.8090	0.8090	149.1461	0.7495	0.7495	**0.059567**	**Maximum**
2	24	0.1206	0.9296	27.86887	0.1400	0.8895	0.040126	
3	8	0.0402	0.9698	10.44669	0.0525	0.9420	0.027831	
4	2	0.0101	0.9799	5.20747	0.0262	0.9682	0.011713	
5	2	0.0101	0.9899	3.034622	0.0152	0.9834	0.006514	
6	0	0.0000	0.9899	1.952028	0.0098	0.9932	0.003295	
	2	0.0101	1.0000	1.344233	0.0068	1.0000	0.000000	
	199	1.0000		**199**	1.0000			

Table 4 depicts Productivity Trend with proportion of authors for NACLIN where only first authors were considered with exponent value of α= 2.42.

Dmax =0.059567

Dmax = |Fo(x) -Sn(x)| = 0.059567

At 0.01 level of significance, K. S. Static = 1.63/sqrt (199) = 0.115548

Dmax = 0.059567<0.115548, therefore data fits in to generalized form of Lotka's law with exponent value of α= 2.42.

Table 5 Productivity Trend: Proportion of author (IASLIC)

No. of contribution	No. of Authors Observed	Observed	Sn(x)	No. of Authors Expected (An)	Expected	Fo(x)	|Fo(x) - Sn(x)|	
1	308	0.7549	0.7549	308.8677	0.7570	0.7570	0.002100	
2	51	0.1250	0.8799	55.3628	0.1357	0.8927	0.012793	**Maximum**
3	17	0.0417	0.9216	20.25404	0.0496	0.9423	0.020769	
4	18	0.0441	0.9657	9.923473	0.0243	0.9667	0.000973	
5	7	0.0172	0.9828	5.705935	0.0140	0.9806	0.002198	
6	4	0.0098	0.9926	3.630424	0.0089	0.9895	0.003104	
7	2	0.0049	0.9975	2.477019	0.0061	0.9956	0.001935	
8	1	0.0025	1.0000	1.778727	0.0044	1.0000	0.000026	
	408	1.0000		408	1.0000			

Table 5 depicts Productivity Trend with proportion of authors for IASLIC where only first authors were considered with exponent value of α= 2.48.

Dmax = 0.012793

Dmax = |Fo(x) -Sn(x)| = 0.012793

At 0.01 level of significance, K. S. Static = 1.63/sqrt (620) = 0.080697

Dmax = 0.012793<0.080697, therefore data fits in to generalized form of Lotka's law with exponent value of α= 2.48.

The analysis of the data revealed that the maximum difference (Dmax) in observed and estimated cumulative proportions is less than the K. S. static (Critical Value) at 0.01 level of significance for all sets of data. Therefore present set of data fits in to generalize form of Lotka's law.

6. Conclusion:

The productivity distribution does not fit Lotka's inverse square law applied to it, in overall as well as individual data sets of CALIBER, NACLIN & IASLIC. However the Lotka's law fits in to set of data by applying K-S test. Productivity Trend with proportion of authors for overall proceedings where only first authors were considered with exponent value of α= 2.15, 2.77, 2.42 and 2.48 for overall, CALIBER, NACLIN and IASLIC proceedings respectively. The analysis of the data revealed that the maximum difference (Dmax) in observed and estimated cumulative proportions is less than the K-S. Static (Critical Value) at 0.01 level of significance for all sets of data as well as CALIBER, NACLIN and IASLIC data. Therefore present

set of data fits in to generalize form of Lotka's law.

References:

Kulkarni, J. N. (2011). Mapping CALIBER, NACLIN & IASLIC Proceedings: A Scientometric Study, Unpublished Doctoral Thesis, Dr. BabasahebAmbedkarMarathawada University, Aurangabad.

Sangam, S.L. (2008). Areas of research in the field of Scientometrics and Informetrics. In M.M. Koganuramath, B.D Kumbar, & B.S. KademaniIn*Professor C.R. Karisiddappa festschrift: Library and Information Science profession in the knowledge society* (pp.265-272). New Delhi: Allied publishers Pvt. Ltd.

Vaishnav, A A andAghav, U R (1994). Authorship pattern and authors productivity. In Rawat, P P (Ed). Current development in library and information science (Kaula festschrift volume). 477-488.