DESIDOC Journal of Library & Information Technology, Vol. 38, No. 4, July 2018, pp. 266-270, DOI : 10.14429/djlit.38.4.12361

Lotka's Applicability on Global Dengue Research Publication : A Scientometric Study

Sadik Batcha M

Annamalai University, Annamalainagar, Tamil Nadu-608 002, India
E-mail: msbau@rediffmail.com

ABSTRACT

Dengue is currently regarded globally as the most important mosquito-borne viral disease. The true impact of dengue globally is difficult to ascertain due to factors such as inadequate disease surveillance, misdiagnosis, and low levels of reporting. Currently available data likely grossly underestimates the social, economic, and disease burden. Estimates of the global incidence of dengue infections per year have ranged between 50 million and 200 million, there are about 6,031 research paper collectively contributed by 22,234 scientist published in 6007 scientific periodicals in recent 8 years as the focus on this fever attracts attention of Scientists at the considerable rate. The present paper attempts to check the applicability of Lotka's Law on the global publication on Dengue. The study lights on Lotka's empirical law of scientific productivity, i.e., Inverse Square Law, to measure the scientific productivity of authors, to test Lotka's Exponent value and the K.S test for the fitness of Lotka's Law.

Keywords: Scientometrics; Lotka's law; Lotka's empirical law; Lotka's exponent value; Kolmogorov-Smirnov test; Dengue research output

1. INTRODUCTION

The frequency distribution of scientific productivity[1], pertaining to scattering of literature[2], and concerning word frequency in the text[3] are the three basic Laws of Bibliometrics. Alfred J. Lotka in his paper on "Analysis of the number of publications in Chemical Abstracts from 1907 to 1916 on frequency distribution of scientific productivity" proposed an inverse square law of scientific productivity. These laws, extensively studied by Fairthorne, are found to show some stable regularities[4]. In general it is inferred that most names occur few times, and a few names occur many times in a study. Lotka in his classic paper published on frequency distribution of scientific productivity presented an analysis of the number of publications listed in Chemical Abstracts from 1907 to 1916 with the frequency of publications of the authors and proposed an inverse square law of scientific productivity[5]. 'if X authors contribute exactly one paper each, then the number of the authors contributing n papers will be expressed in the terms of the equation

$A_n = X/n^2$ for $n = 1, 2, 3, \ldots$etc

where A_n= number of authors contributing n papers; x=number of authors contributing one paper; n = 1,2,3,.....i. It is also expressed as $1/n^2$ where $1= X$

Lotka's article was not cited for 15 year until 1941 and his distribution was not termed as Lotka Law until 1949. This law has been tested by many studies. Murphy[6] mapped in humanities and Schorr[7] on librarianship, Pao[8] has also tested this law. Booksteen[9] gave modified theoretical models for Lotka's Law

$$Y_a = k/x^a \quad (1)$$

where Yx = the relative frequency of authors publishing

x = the number of papers

k & $á$ = constant

Considering the fact that n authors have produced one article each, the value of constant k can be easily derived, by putting the value of $Y_a = 1$ and $a = 2$.

$Y_a = k/x^a$; $1 = k/1^2$; $K = 1$

In the other words the number of authors with x papers is proportional to $1/x^2$.

This law was statistically verified by Murphy in the field of humanities and by Schorr in the area of map literature. Bookstein further studied Lotka's Law in detail and generalised the equation as:

$an = kx/ná$ for $n = 1,2,3\ldots$..

The degree of conformity or non-conformity of various empirical distributions to Lotka's distribution has been tested by many prominent authors. The studies on the fitness of Lotka's Law began in a systematic manner that started with the work of Pao using least square method with 48 sets of authors' productivity data. The study found that majority of data sets conform to Lotka's distribution as a general inverse power function (á = 2). Lemoline's study[10] regarding the productivity pattern of the CSIR scientists revealed that the distribution corresponding to the entire authors population does not follow an inverse power relationship – either general or square and showed the existence of a flat distribution amongst those

Received : 24 November 2017, Revised : 18 April 2018
Accepted : 26 April 2018, Online published : 25 June 2018

researchers with 10 or less number of articles and of an inverse square power relationship for those who have written more than 10 article. Radhakrishnan and Kernizan[11] concluded that Lotka's Law in the generalised form seems applicable while considering the publication of authors in one periodical whereas when considering the publications of authors in various journals, the observed values deviate considerably from the value predicted by the law. Findings of Kumar[12], Amsaveni[13] and Batcha[14] corroborate with the Lotka's findings and values of 'á' were found to be around two. The study on Dengue by Dwivedi[15] reveals that Number of publications increases 70 fold in given time span during 1989-2015. Vietnam has the highest impact (quality) while India has the lowest impact among the leading countries. The study of Aloe Vera research by Gupta[16] revealed that Dengue publications registered 3.9 per cent growth and averaged its citation impact to 9.57 citation per paper. The top 12 most productive countries individually contributed global share 5.36 per cent to 52.17 per cent.

Dengue is a mosquito-borne viral disease that has rapidly spread in all regions of WHO in recent years. Severe dengue (also known as Dengue Haemorrhagic Fever) was first recognised in the 1950s during dengue epidemics in the Philippines and Thailand. Today, severe dengue affects most Asian and Latin American countries and has become a leading cause of hospitalisation and death among children and adults in these regions. The incidence of dengue has grown dramatically around the world in recent decades. The actual numbers of dengue cases are underreported and many cases are misclassified. One recent estimate indicates 390 million dengue infections per year (95 % credible interval 284 million –528 million), of which 96 million (67 million – 136 million) manifest clinically (with any severity of disease)[17]. Another study, of the prevalence of dengue, estimates that 3.9 billion people, in 128 country, are at risk of infection with dengue viruses.

2. OBJECTIVES

This paper has following objectives:

(a) To analyse the author productivity patterns in the field of Global Dengue literature

(b) To examine the validity of Lotka's law using total counting and straight counting of authors

(c) To apply Kolmolgorov- Smirnov (K-S) goodness-of-fit test for the conformity of Lotka's law.

3. METHODOLOGY

The study retrieved and downloaded the publication data of the world and of most productive authors on Dengue Research from the Web of Science database during 2010-17. (till October 2017) on 05-11-2017. The study retrieved and downloaded the publication data of the world and of the most productive authors on Dengue Research from the Web of Science database during 2010-17. The Search word 'Dengue Disease' and 'Dengue Fever' was used in 'title', abstract and keyword" search tag and restricted it to the period 2010-17 in 'Period range' search tag for searching global publication data. The searches were performed on the names of Dengue Disease and Dengue fever with all probabilities and bibliographical details of 6,031 research paper collectively contributed by 22,234 scientist published in 6007 scientific periodical were collected for application of Lotka's Law.

4. DATA ANALYSIS AND DISCUSSION

It is observed from the data for the period 2010 to 2017 (Table 1) that 6031 research Publication were brought out with collective responsibility as well as sole authorship. 288 author have one, 656 author have two, 769 author have three and 735 author have four papers each to their credit and so on. The maximum number of papers by an individual is found to be 288 which reflect 4.78 per cent of total publications at global level. Tests have been conducted to find out as to what extent, the author productivity conforms to the Lotka's Law.

Lotka, based proposed an empirical law of scientific productivity, i.e., Inverse Square Law, to measure the scientific productivity of authors as:

$$Yx = k/x2 \qquad (1)$$

Bookstein presented a generalised form of Lotka's Inverse Square Law as Lotka's Inverse Power Law as:

$$Yx = k/x\acute{a} \qquad (2)$$

where Yx = the relative frequency of authors publishing

x = the number of papers

k & $\acute{a}$ = constant

The Lotka's law and its modified version have a wide range of applicability to a variety of phenomena and the mere form of such a distribution throws a little or no light on the underlying physical conditions. The fact that exponent has approximately the value of 2 as depicted, enables us to state the results in the following simple form:

$$Yx = 6/ð2x2 \qquad (3)$$

as, $k = 6/ð2$ for á =2

The formula has two part: an exponent with value 2 in the inverse square, i.e., *1/x2* portion of expression and a constant 6/ð2 (or 600/ð2 in percent) determined by the value of the exponent. Considering the fact that the data in the present study 16278 author have produced one article each, the values of constant k can be easily derived, by putting the value of Yx, i.e., 1 and á =2 in Eqn. (1).

16278 = *k/1á*

By taking the value of á as 2 (Lotka's exponent), the expected values as in Table 2 are obtained.

It can be observed from the Table 2 that the expected values of are not close to the observed values. The frequency distribution of authorship of Dengue research contribution is tested by applying the Lotka's law. An attempt is made to find out the Lokta's exponent value for the Global Scientists output on Dengue publications, calculated by using the formula proposed by Pao[18].

$$\acute{a} = \frac{N\,\acute{O}\,XY - \acute{O}\,X\,\acute{O}\,Y}{N\,\acute{O}X2 - (\acute{O}X)2} \qquad (4)$$

where N = no. of the pairs of data considered

X = the logarithm of x

Y = the logarithm of y

Table 3 presents the values which are determined from the logarithmic table and required for Eqn. (4). Based on Eqn. (4), the Lotka's exponent is calculated for Global Research

Table 1. Year wise dengue research publication output with authorship pattern

Year	One	Two	Three	Four	Five	Six	Seven	Eight	Nine	Ten & above	Total	Percent
2010	35	58	67	74	62	53	51	33	21	91	545	9.04
2011	27	64	53	63	68	67	39	38	34	91	544	9.02
2012	46	60	80	74	71	63	56	39	36	109	634	10.51
2013	24	82	100	85	84	59	66	57	33	128	718	11.91
2014	30	87	96	89	89	71	66	65	42	129	764	12.67
2015	40	114	113	107	98	92	71	41	46	163	885	14.67
2016	55	95	151	142	115	112	80	78	74	205	1107	18.36
2017	31	96	109	101	90	85	75	46	38	163	834	13.83
Total	288	656	769	735	677	602	504	397	324	1079	6031	100.00
%	4.78	10.88	12.75	12.19	11.23	9.98	8.36	6.58	5.37	17.89	100.00	

Table 2. Number of expected authors derived with the value of å = 2

No of papers	Obs. No of authors	%	Total authorship	Expected no of authors	Percentage	Difff in authors	Diff in %
One	16278	100.00	16278	16278	100.00	0	0.00
Two	3045	18.71	6090	4070	25.00	1025	6.29
Three	1142	7.02	3426	1809	11.11	667	4.10
Four	625	3.84	2500	1017	6.25	392	2.41
Five	331	2.03	1655	651	4.00	320	1.97
Six	221	1.36	1326	452	2.78	231	1.42
Seven	159	0.98	1113	332	2.04	173	1.06
Eight	102	0.63	816	254	1.56	152	0.94
Nine	73	0.45	657	201	1.23	128	0.79
Ten & Above	258	1.58	2580	163	1.00	-95	-0.58
Total	22234		36441	25227			

Table 3. Frequency Distribution for Lotka's Exponent of Dengue Research Publication

X	*Y*	*X=Log X*	*Y=Log Y*	*XY*	*X2*
1	16278	0.00000	4.211601	0.000000	0.000000
2	3045	0.30103	3.483587	1.048664	0.090619
3	1142	0.47712	3.057666	1.458877	0.227645
4	625	0.60206	2.795880	1.683287	0.362476
5	331	0.69897	2.519828	1.761284	0.488559
6	221	0.77815	2.344392	1.824292	0.605519
7	159	0.84510	2.201397	1.860396	0.714191
8	102	0.90309	2.008600	1.813947	0.815572
9	73	0.95424	1.863323	1.778062	0.910579
10	41	1.00000	1.612784	1.612784	1.000000
		6.559763	26.099059	14.841594	5.215159

Publication on Dengue as á = 2.498,

$$n = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - \left(\sum X\right)^2} = \frac{10(14.841594) - (6.559763)(26.099059)}{10(5.215159) - (6.559763)^2} = -2.498$$

while theoretical Lotka's value is á = 2.000.

Theoretical Value of *n* Value 2.498 is matched with the table value of R.ROUSSEAU for getting C.S value -0.7454.

Constant value of present study	*n* value
0.7454	2.498
Lotka's constant value	*n* value
0.6079	2

D-Max value of present study = 0.0064

D-Max value of Lotka's study = 0.1314

Fe+ = $0.7454(1/X^{2.498})$ Fe+ = $0.6079(1/X^{2.000})$ D-Max = 0.0064

Critical value at .01 level of significance = $2.498/\sqrt{22017} = 0.0168$

The theoretical value of C as 0.7454 is taken from Table No. 4.6.6. in the book "Introduction to Informetrics[19]". To apply the Kolmogorov - Smirnov test for the fitness of Lotka's Law to global Research Publications on Dengue, Tables 4 presents the expected value of distribution of authors in research publications and theoretical value of Lotka to find out the D-Max values.

5. RESULTS AND FINDINGS

Generally Lotka's Law describes the frequency of publications by authors in a given subject/discipline. In this paper, an attempt has been made to study the applicability of

Table 4. K.S Test of observed and expected distribution of authors in dengue research publication

x	y_x	Observed = $yx/\sum yx$	Value = $\sum(yx/\sum yx)$	Expected Freq.	Value of Freq./Cum.	Diff. (D)	Expected Freq.	Value of Freq./Cum.	Diff. (D)
1	16278	0.7393	0.7393	0.7454	0.7454	-0.0061	0.6079	0.6079	0.1314
2	3045	0.1383	0.8776	0.1320	0.8774	0.0064	0.1520	0.7599	-0.0137
3	1142	0.0519	0.9295	0.0479	0.9253	0.0039	0.0675	0.8274	-0.0157
4	625	0.0284	0.9579	0.0234	0.9486	0.0050	0.0380	0.8654	-0.0096
5	331	0.0150	0.9729	0.0134	0.9620	0.0017	0.0243	0.8897	-0.0093
6	221	0.0100	0.9829	0.0085	0.9705	0.0016	0.0169	0.9066	-0.0068
7	159	0.0072	0.9902	0.0058	0.9763	0.0014	0.0124	0.9190	-0.0052
8	102	0.0046	0.9948	0.0041	0.9804	0.0005	0.0095	0.9285	-0.0049
9	73	0.0033	0.9981	0.0031	0.9835	0.0002	0.0075	0.9360	-0.0042
10	41	0.0019	1.0000	0.0024	0.9858	-0.0005	0.0061	0.9421	-0.0042

the Lotka's Law to the Global publications on a subject or a discipline of Dengue fever Disease.

A K-S test is applied for the fitness of Lotka's law fits to the global Dengue data. Result indicates that the value of D-Max, i.e, 0.0064 determined with Lotka's exponent, á = 2.498 and is not close to the D-Max value 0.1314 determined with the Lotka's exponent á = 2 than the critical value determined at the 0.01 level of significance, i.e., 0.0168. Thus, distribution frequency of the authorship follows the exact Lotka's Inverse Law with the exponent á = 2. However, the modified form of the inverse square law, i.e. Inverse Power Law with á and C parameters as 2.498 and 0.7454 is not applicable and does not appear to provide a good fit. Lotka's law of author productivity is regarded as one of the classical laws of bibliometric.

The present study showed that Lotka' generalised law is not applicable to Dengue literature. K-S test is applied to verify the applicability of Lotka's law of scientific productivity. The statistical tests show that the Lotka's law in its generalised form does not fit the author productivity distribution pattern prepared for the straight count and for the contribution of complete count of the Dengue Research Output.

6. CONCLUSIONS

The results obtained in this study do not follow the inverse square power law of Lotka as such and similarly. B.K. sen[20] in his short communication regarding Narendra Kumar's[21] Article stated that the Value of á as method. The difference between two values is rather small. i.e. 0.16. Both the cases the values are quite close to actual values indicating that the data set 2 of NK by and large follows Lotka's Law. The works of Huber[21-23] describe the statistical problems related to the fitting of the Lotka's formulation and related regularities. Nicholls[24] has analysed about 15 studies conducted during 1973 and 1986, and observed that the result of these studies do not provide the clear-cut validation of the law.

REFERENCES

1. Lotka, A. J. The frequency distribution of scientific productivity, *J. Wash. Acad. Sci.*, 1926, **16,** 317-323.
2. Bradford S.C. Documentation. D.C, Public Affairs Press, Washington, 1950.
3. Zipf, G. Human Behaviour and the principle of least effort. Addison-Wesley, Cambridge, (MA), 1949.
4. Fairthorne, R. A. Progress in documentation: Empirical hyperbolic distributions (Bradford-Zipf-Mandelbrot) for Bibliometric description and prediction, *J. Doc.*, 1969, **25**, 319.
5. Lotka, A.J. The frequency distribution of scientific productivity, *J. Wash. Acad. Sci.*, 1926, **16**, 317-323.
6. Murphy, L.J. Lotka's Law in the humanities. *J. Am. Soc. Inf. Sci.*, 1973, **24**(6), 461-462.
7. Schorr, A.E. Lotka's Law and map librarianship. *J. Am. Soc. Inf. Sci.*, 1975, **26** (2), 189-190.
8. Pao, M.L. An empirical examination of Lotka's Law. *J. Am. Soc. Inf. Sci.,* 1986, **37**, 26-33.
9. Bookstein, A. Pattern of scientific productivity and social change: A discussion of Lotka's Law and bibliometric symmetry. *J. Am. Soc. Inf. Sci.*, 1977 **28**(4), 206-210.
10. Lemoline, W. The frequency distribution of research papers and patents according to sex: The cases of CSIR, India. *Scientometrics*, 1992, **24**(3), 449-469.
11. Radhakrishnan, T. & Kernizen, R. Lotka's Law and computer science literature. *J. Am. Soc. Inf. Sci.*, 1979, **30**, 51-54.
12. Narendra Kumar. Applicability to Lotka's law to research productivity of Council of Scientific and Industrial Research (CSIR), India. *Annals Lib. Inf. Stud.*, 2010, **57,** 1-5.
13. Amsaveni, N. & Batcha, Sadik. M. Bibliometric dimension of gender studies in informatics from G-8 countries. *Ind. J. Inf. Sci. Serv.*, 2009, **1**(3), 65-71.
14. Batcha, Sadik. M. Research output analysis on robotic technology: A Scientometric study. *Ind. J. Inf. Sci. Serv.*, 2017, **7** (1), 25-31.
15. Sandhya Dwivedi. Dengue research: Three dimensional bibliometric study of the global research output during 1989-2015. *DESIDOC J. Lib. Inf. Technol.*, 2017, **37**(3), 180-185.
16. Gupta, B.M.; Ahmed, K.K.M.; Dhawan, S.M.; Gupta, R. *Aloe Vera* (Medicinal Plant) research: A scientometric assessment of global publications output during 2007-16. *Pharmacog. J.* 2018, **10**(1), 1-8.

17. Bhatt, S.; Gething, P.W.; Brady, O.J.; Messina, J.P.; Farlow, A.W.; Moyes, C.L.; Drake, John M.; Brownstein, John S.; Hoen, Anne G.; Sankoh, Osman; Myers, Monica F.; George, Dylan B.; Jaenisch, Thomas; William, Wint G.R.; Simmons, Cameron P.; Scott, Thomas W.; Farrar, Jeremy J. & Hay, Simon I. The global distribution and burden of Dengue, *Nature*, 2013, **496**, 504-507.
18. Pao, M.L. Lotka's law: A testing procedure. *Inf. Process. Manag.*, 1985, **21**(4), 305-320.
19. Egghe, L. & Rousseau, R. Introduction to informetrics: Quantitative methods in library. Documentation and Information Science. Elsevier Science Publishers, Amsterdam. 1990, 357.
20. Sen, B.K. Lotka's Law: A view point. *Annals Lib. Inf. Stud.*, 2010, **57**, 166-167.
21. Kumar, Narendra. Applicability to Lotka's law to research productivity of council of scientific and industrial research CSIR, India, *Annals Libr. Inf. Stud.*, 2010, **57**, 1-5.
22. Huber, J.C. A new model for analysing scientific productivity, *J. Am. Soc. Inf. Sci. Technol.*, 2001, **52**, 1089-1099.
23. Huber, J.C. A new model that generates Lotka's law. *J. Am. Soc. Inf. Sci. Technol.*, 2002, **53**, 209–219.
24. Nicholls, P.T. Bibliometric modeling processes and the empirical validity of Lotka's law. *J. Am. Soc. Inf. Sci.*, 1989, **40**(6), 379-85.

CONTRIBUTOR

Dr. M. Sadik Batcha, has done his MLIS, MPhil and PhD from Annamalai University. Presently working as the Associate Professor in Deprtment of Library and Information Science, Annamalai University. His area of research includes Scieentometrics, Altmetrics. He has published 60 research papers in journals. He has edited 5 book and has been a life member of ALA.