

Lotka's Law a and The Pattern of Scientific Productivity in the Gandhian Literature

Shiv Singh

Madan Singh Rana

Abstract

This paper offers a practical insight into the application of Lotka's law of author productivity in Gandhian literature and also presents the collaborative pattern among the authors. In the present work significant effort has been made to collect all the periodical literature related to Gandhi and that are published in English up to 2010 from the various resources. The study reveals many imperative indicators concerning Gandhian literature, like for instance that Lotka's Law holds with respect to Gandhian Literature.

Keywords: Bibliometrics, Authorship Pattern, Lotka's Law, Scientific Productivity, M. K. Gandhi, Gandhian Literature

1. Introduction

All information, which is published in the form of various documents for any discipline, is only due to the thought of somebody who is called an 'Author'. An author is one who brings something into being, or in other words, he/she is a beginner of any action or state of that thing. Authors are of prime importance as they not only produce, new scientific knowledge in different fields, but, are also concerned with the propagation and the use of information. Hence, it is essential to investigate the role of the author in particular field i.e., authorship pattern. This study determines the extent of citations received by different authors. It also provides the information related to the individual author's productivity, which is very useful in determining the status of the author among his/her colleague within a field. Besides this, authorship patterns deal with the productivity of authors. Thus in general, it is a study related to authors, their productivity, collaborative patterns and other aspects.

Laws have been developed to calculate the frequency and distribution of scientific productivity. Among them is Lotka's law¹ given by Alfred Lotka in 1926. It describes the frequency of publication by authors in a given field. It states 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 in the region of 60 per cent". This implies that among all the authors in a given field, 60 per cent will have single publication; 15 per cent will have two publications i.e. $1/2^2$ times 0.60; 7 per cent will have three publications ($1/3^2$ times 0.60), and so on. Hence, Lotka's proposed the following inverse square law of scientific productivity.

$$a_n = c/n^b, \quad (1)$$

where $c > 0, \quad b > 1$

- c - is a constant
- n - 1,2,3, n
- b - characteristic exponent of the distribution for a set of data and
- a_n - number of authors contributing n number of paper each.



Lotka's equation depends on three factors :

- (i) The number of contributors with minimal productivity (authors with single paper each);
- (ii) The maximum productivity of contributors (n_{max}); and
- (iii) The characteristic exponent of the distribution for a set of data.

Productivity trends for well-established and recognized subjects and disciplines follow distribution patterns which confirm Lotka's law if applied to such distribution in its original form with an exponent value of two. Collaboration is a major area of study, which indicates how authors work together to bring out a publication. Scientific collaboration is usually conceptualized as works of a team of scientists with shared goals, such as formulating or testing particular empirical hypothesis, and with shared products, such as co-authored papers.

2. Literature Review

Extensive research has been conducted on the validation of Lotka's law²⁻¹¹. However, most of the studies have verified the law and they repeatedly found that the exponent n is not always 2 but it is rather a variable. Likewise, Pao² considered datasets of 48 author covering 20 subject fields and 3 huge research library catalogues. She predicted that more than 80% of the datasets obey Lotka's law and the value of $n=2$ for only seven datasets. This kind of results are being reported by the other researchers as well.³⁻⁵

Using maximum likelihood approach Nicholls⁶ applied Lotka's law and estimated the various parameters. The results obtained by it are comparatively better. The maximum likelihood method and the least squares methods have been compared by Newman⁷. He reported that least squares method

generally overestimates the results because of statistical fluctuations in logarithm of data. Newby et al.⁸ tested the productivity of the programmers in open source software development using the Lotka's law. He considered programmers as authors and software as publication. The found exponent $n=2.82$.

Petek⁵ investigated the personal name headings in the Slovenian online catalogue COBIB using Pao's method. He reported the value of $n = 2.2656$ and constant $c = 0.6890$. He concluded that Lotka's law holds for the occurrence of personal name headings in COBIB. Ahmed and Rahman⁹ examined the validity of Lotka's law in the fields of the authors of nutrition research in Bangladesh. They have used least squares method to check the applicability of Lotka's law

3. Scope and Objectives

In the field of Gandhian Studies, the available literature is widely scattered. Most of Gandhian literature is published by the scholars' time-to-time to measure its value based on different criteria. So access to appropriate and relevant literature by researchers becomes both critical and difficult. Similarly, the application of bibliometric indicators in the field of Gandhian literature has not been explored yet. Thus, in present work efforts have been made to develop a database of the available literature and apply bibliometric indicators to know the authorship pattern and scientific productivity in Gandhian literature. Basically, the present study is a part of research entitled 'Periodical literature in Gandhian Studies: A bibliometric Study' and is confined to the periodical literature published in English up to year 2010. For this purpose, a significant effort has been made to collect all the periodical literature from the various primary sources like Asian Survey, Economic

and Political Weekly, Gandhi Marg, Gandhian Perspectives, Journal of Peace Research, Mainstream, Philosophy East and West, Social Scientist, The Journal of Asian Studies, and Journal of Peace Studies. In addition to these, every possible effort has also been made to include Gandhian literature by scanning the secondary resources for a better coverage. Likewise J-STOR, Project Muse, Science Direct, Proquest, Psycinfo, LISA, Eric, Pilots, Sociological Abstracts, Social Services Abstract Index, International Political Science Abstracts, Social Science Citation, Guide to Indian Periodical Literature, J-GATE, Econlit, Scopus, Web of Science, Emerald Insight, Cambridge University Press, Oxford University Press, Wiley-Blackwell, Taylor & Francis, Springer, Sages, ACM Portal, IEEE explore etc.

The present study is endeavor to find the authorship pattern and scientific productivity in the Gandhian literature in terms of Contribution of authors in Gandhian literature and application of Lotka's law.

4. Methodology

In the present work, significant effort has been made to collect all the periodical literature related to Gandhi published in English up to 2010 from the above mentioned resources. Every article has been

scanned and all necessary bibliographic information was entered in the Excel for analysis. This resulted in 6175 articles.

5. Results and Discussions

5.1 Applicability of Lotka's Law

In order to apply the Lotka's law the author index was generated from the computerized database on Gandhian literature. The frequency of authors and number of contributions were calculated and placed against it respectively. The testing procedure proposed by Pao in 1985 has been used for fitting of Lotka's inverse – power rule. The basic steps followed were:

- 1) Table 1 was created in which a frequency consisting of pairs (n, an); here 'n' represents the number of contributions, and 'an' is the number of authors writing 'n' paper each.
- 2) Adopted Lotka's inverse-power law, as given in equation (1).
- 3) The parameter 'b' was estimated by the least-square method in the simple regression model:

$$\text{Log } a_n = \log c/n^b$$

$$\text{Log } a_n = \log n^{-b} + \log c$$

$$\text{Log } a_n = \log c - b \log n \quad n = 1, 2, 3, \dots \dots \dots (2)$$

Table 1:- Contribution of Authors

No of Contribution	No of Authors	% of Author	Total Contribution	% of Contribution
1	2826	74.31	2826	44.82
2	503	13.23	1006	15.96
3	193	5.07	579	9.18
4	90	2.37	360	5.71
5	51	1.34	255	4.04
6	36	0.95	216	3.43

7	26	0.68	182	2.89
8	21	0.55	168	2.66
9	11	0.29	99	1.57
10	12	0.32	120	1.90
11	14	0.37	154	2.44
12	7	0.18	84	1.33
13	2	0.05	26	0.41
15	2	0.05	30	0.48
17	2	0.05	34	0.54
18	1	0.03	18	0.29
19	2	0.05	38	0.60
20	1	0.03	20	0.32
23	1	0.03	23	0.36
27	1	0.03	27	0.43
40	1	0.03	40	0.63
Total	3803	100.00	6305	100.00

Table 2 gives the value of 'b' and estimated frequency of author by using equation (2). The average value of 'b' was found to be 2.48. By using this value, the estimated frequencies of authors were calculated and are also presented.

- 4) The parameter 'c' was equal to number of contributors with minimal productivity (i.e. 2826 number of authors)
- 5) Lastly, the Kolmogorov-Smirnov (K-S) goodness-of-fit test was applied to ascertain whether the data fits the model. The K-S test used to test the maximum deviation (D) is: $D_{max} = |F_o(x) - S_n(x)|$ where 'F_o(x)' is the observed relative frequency of a sample of total number of authors (n=3803). Since the K-S D_{max} of 0.0016 has been found greater than the level of significance (p=0.01) of 0.0264, as (Table 2), the data conforms to Lotka's distribution. Figure 1., illus-

trates the variation of observed and estimated authors with their contributions.

Average Value of $b = 2.48$

Thus, on applying Lotka's law we get following parameters:

- (i) Total number of authors, $n=3803$
- (ii) $D_{max} = F_o(x) - S_n(x) = 0.0016$
- (iii) Critical Value at the 0.01 level of significance
- (iv) K-S Statistics = $1.63/\sqrt{n} = 1.63/\sqrt{3803} = 1.63/61.66$
- (v) $D < 0.0264$

Since, $D < 0.0264$, therefore, data fit Lotka's generalized formula with exponent value (b) of 2.48.

Table 2:- Frequency of Estimated Authors

No of Contribution	Value of b	Estimated Frequency of Authors	% of Authors
1		2826	74.63
2	2.49	507	13.38
3	2.44	185	4.89
4	2.49	91	2.40
5	2.49	52	1.38
6	2.44	33	0.88
7	2.41	23	0.60
8	2.36	16	0.43
9	2.53	12	0.32
10	2.37	9	0.25
11	2.21	7	0.20
12	2.41	6	0.16
13	2.83	5	0.13
15	2.68	3	0.09
17	2.56	3	0.07
18	2.75	2	0.06
19	2.46	2	0.05
20	2.65	2	0.04
23	2.53	1	0.03
27	2.41	1	0.02
40	2.15	0	0.01
Total	2.48	3787	100.00

Table 3:- Frequency of Observed and Estimated Number of Authors (Kolmogorov-Smirnov Test)

No of Contribution	Observed		Estimated		Deviation D=Fo(x)-Sn(x)	Dmax
	Cumulative Frequency	Relative Frequency	Cumulative Frequency	Relative Frequency		
		Sn(x)		Fo(x)		
1	2826	0.7431	2826	0.7462	-0.0031	0.0016
2	3329	0.8754	3333	0.8800	-0.0046	
3	3522	0.9261	3518	0.9289	-0.0028	
4	3612	0.9498	3609	0.9529	-0.0031	

5	3663	0.9632	3661	0.9667	-0.0035	
6	3699	0.9727	3694	0.9755	-0.0028	
7	3725	0.9795	3717	0.9814	-0.0020	
8	3746	0.9850	3733	0.9857	-0.0007	
9	3757	0.9879	3745	0.9890	-0.0010	
10	3769	0.9911	3755	0.9914	-0.0004	
11	3783	0.9947	3762	0.9934	0.0014	
12	3790	0.9966	3768	0.9949	0.0016	
13	3792	0.9971	3773	0.9962	0.0009	
15	3794	0.9976	3776	0.9971	0.0005	
17	3796	0.9982	3779	0.9978	0.0004	
18	3797	0.9984	3781	0.9984	0.0000	
19	3799	0.9989	3783	0.9989	0.0001	
20	3800	0.9992	3784	0.9993	-0.0001	
23	3801	0.9995	3786	0.9996	-0.0002	
27	3802	0.9997	3786	0.9998	-0.0001	
40	3803	1.0000	3787	0.9999	0.0001	

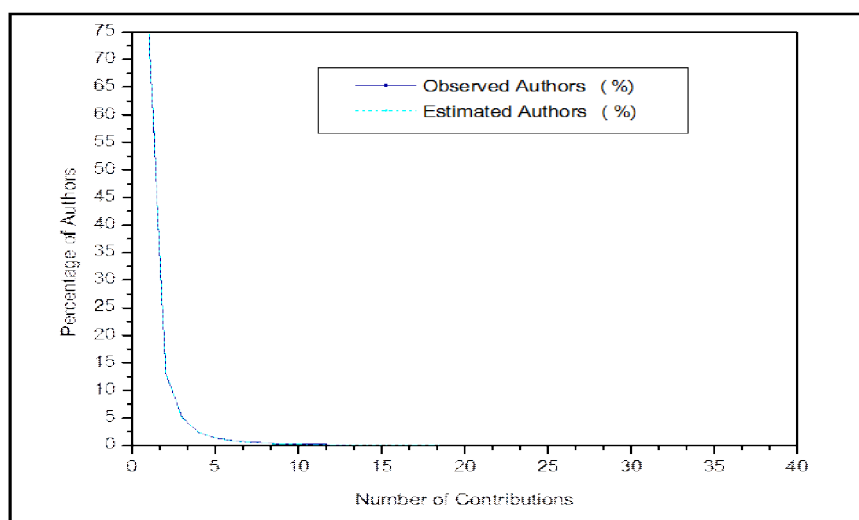


Figure 1:- Trend of observed and estimated authors with their contributions

5.2 Collaborative Pattern in Gandhian Literature

In order to investigate the collaboration among the authors in Gandhian literature for period under con-

sideration, the authorship pattern was analyzed to determine the percentage of single author, two, and multiple authors. In this paper, out of 6175 articles, 5936 articles were considered as 239 articles were

without authors. These articles were contributed by 3803 unique authors. Thus overall average of authors comes to be 1.56 articles per year. Table 4 reflects that single authored papers were the maximum with 95.22 %. On the other hand, contribution

by two authors accounted for 4.06 % and for three authors was 0.42 %. The analysis of the data shows that the maximum number of authors who have contributed to the journal had a tendency to work as individual.

Table 4:- Number of Authors in Gandhian Literature

Authorship	No. of Contribution	% of Contribution	Cumulative %
Single Author	5652	95.22	95.22
Two Author	241	4.06	99.28
Three Author	25	0.42	99.70
Four Author	7	0.12	99.81
Five Author	6	0.10	99.92
Six Author	3	0.05	99.97
Seven Author	1	0.02	99.98
Thirteen Author	1	0.02	100.00

5.2.1 Single vs Multiple Authorship Trend in Gandhian Literature

The decade-wise trend in single vs multiple authorship has been summarized in Table 5. It was clear that for period under consideration single authored papers were maximum i.e. 5652 (95.22 %). It is worth mentioning that the growth of literature has in-

creased decade wise continuously by the contributions of individual authors. Multiple authored papers have shown increasing trend from 4.23 % (1971-1980) to 7.46 % (2001-2010), however their contribution is very less in comparison to single author. Moreover, it can be said that, researchers are trying to work in team spirit.

Table 5:- Single vs Multiple Authorship Trend in Gandhian Literature

Year	With Single Author	With Multiple Author	Total
1920	1 (100%)	0 (0%)	1
1921-1930	24 (96%)	1 (4%)	25
1931-1940	25 (100%)	0 (0%)	25
1941-1950	67 (97.1%)	2 (2.9%)	69
1951-1960	261 (99.24%)	2 (0.76%)	263
1961-1970	910 (98.48%)	14 (1.52%)	924
1971-1980	702 (95.77%)	31 (4.23%)	733
1981-1990	884 (95.57%)	41 (4.43%)	925
1991-2000	1340 (94.57%)	77 (5.43%)	1417
2001-2010	1438 (92.54%)	116 (7.46%)	1554
Total	5652 (95.22%)	284 (4.78%)	5936

6. Conclusion and Recommendation

On the basis of analysis of authorship pattern and collaborative pattern in Gandhian literature, the following deductions have been drawn:

Application of Lotka's Law in Gandhian Literature

It was found that each of 74.31 % of authors contributed only one contribution. The index constant value of authors' productivity has been observed to be 2.48, which is close to the inverse square value (i.e. 2) of Lotka's law. The maximum deviation (D) using Kolmogorov-Smirnov statistics test was found to be 0.0016, at the 0.01 level of significance this value being lesser than 0.0264, upholds the applicability of Lotka's Law to the data of this study. Hence, in simplified way it was found that if the frequency of the authors increases, the number of contributions decreases, whereas if the frequency of the authors decreases, the number of contribution increases.

Collaborative Pattern in Gandhian Literature: Single Vs Multiple Authorship Trend

Most of the contributions were from single authors, as single authored papers were the maximum with 95.22 %. On the other hand contribution by two authors accounted for 4.06 % and for contributions by three authors was 0.42 %. The overall average of authors was 1.56 articles per year. Thus, supporting the fact that authors in social science mostly work as individuals.

The present study is among the first to investigate the authorship and collaborative pattern in the emerging discipline i.e. Gandhian literature. The study has revealed many important parameters and characteristics of Gandhian literature. Further, it also helps to compare with the established disciplines and also perceives the practical utility for library

professionals and Gandhian Admirers/followers/scholars. But still, in the view of forgoing, it is recommended that

- (i) The work can be extended to present year
- (ii) More resources like books can also be explored.

References

1. Lotka, A. J. (1926) The frequency distribution of scientific productivity. *Journal of the Wasington Academy of Science*, vol. 16(12), 317-323.
2. Pao, M. L. (1985). Lokta'slaw : a testing procedure. *Information proceeding management* 21(4) 305-320
3. Patra, S. K. and Mishra, S. (2006) Bibliometric study of bioinformatics literature. *Scientometrics*, 67(3), 477-489.
4. Patra, S. K. and Chand, P. (2007) HIV AIDS research in India: a bibliometric study. *Library & Information Science Research*, 29(1) 124-134.
5. Petek, M. (2008) Personal name headings in COBIE: testing Lotka's law. *Scientometrics*, 75(1), 175-188.
6. Nicholls, P T, (1989) Bibliometric modelling processes and the empirical validity of loktas law. *Journal of American society for information science* 40(6) 379-385.
7. Newman, M. E. J. (2000) Comments to the article by Rousseau and Rousseau. *Cybermetrics*, 4(1), 1-2
8. Newby, G. B. Greenberg, J. and Jones, P. (2004) Open source software development and Lotka's law: bibliometric patterns in programming. *Journal of the American Society*

- for information science and technology, 55(13), 1209-1211.
9. Zayed, S.M. Ahmed and Rahman, Md. Anisur (2009) Lotka's law and authorship distribution in nutrition research in Bangladesh. *Annals of Library and Information Studies* Vol. 56, June. 95-102.
 10. Rana, Madan Singh (2010) *Scientometric study of wild mammal research in India: Authorship, distribution and research*. Germany: Lambert Academic Publishing.
 11. Rajneesh (2012) *Scientometric study of computer science with special reference to literature published by association for computing machinery (ACM)*. Rajasthan: Banasthali University: (Unpublished Ph.D Thesis)

About Authors

Dr. Shiv Singh, Professional Assistant, Learning Resource Center (LRC), Jaypee University, Anoopshahar, District-Bulandshahr, Uttar Pradesh.

Dr. Madan Singh Rana, University Librarian, Central Library, HNB Garhwal University, Srinagar, Garhwal, Uttarakhand,