September 2020

Lotka's Law and Authorship trends in Library and Information Science: A study based on select journals of India, US and UK

GARIMA BISARIA
garimabisaria37@gmail.com

Lotka's Law and Authorship trends in Library and Information Science: A study based on select journals of India, US and UK

Garima Bisaria

Research scholar, Department of Library & Information Science, University of Luck now, Uttar Pradesh

E-mail: garimabisaria37@gmail.com

Abstract

The present study investigates authorship trends and productivity among Library and Information Science (LIS) professionals who contributed journal articles in the selected 12 peer reviewed LIS journals of India, US and UK. For the purpose of the study four periodicals from each country are chosen during the period 2007 to 2017. The results analyzed proved that journal wise author distribution is highest in the published literature of UK with *(3546)* authors followed by India with *(3162)* and US *(2420)* authors. The obtained values of χ^2= 66.331, ρ=0.05 found that there is an association between type of authorship and country. As India and UK accounted for more number of collaborative works while US has large number of solo article contributions. The data analyzed shows preference towards collaborative research in all the three countries with little variations. Lotka's law have been applied and tested using Pao's method and verified through Kolmgorov-Smirnov (KS) test. The application of Lotka's law when applied individually on author's data set of published literature of India with *Dmax value=0.0238* and *critical value=0.046*, proves that observed authorship data holds good for Lotka's law in the authors' data . But *with Dmax value=0.1362 and CV=0.04* in the author distribution of US and *D max value=0.2520* and *CV=0.04* of UK, does not support lotka's law. Overall the results of the K-S tests proved that author productivity distribution does not fits Lotka's law in the subject of library and information science in scholarly research output in Library and information science published journal literature of US and UK.

Keywords: Lotka's Law; publication productivity; author productivity; collaboration coefficient; degree of collaboration; authorship trends;Library and Information Science

Introduction

Alfred J. Lotka (1926) studied author productivity patterns and developed one of the main laws in bibliometrics. He published a classic paper on his study about the frequency distribution of scientific productivity of authors observing the publications listed in Chemical Abstracts for the period 1907-16. He observed that, in a given area of science, there are lot of authors who publish only one study, while a small group of prolific authors contribute with a large number of publications. It is also known as inverse square law on author productivity. The law takes the number of authors who have contributed with a single study and then predicts how many authors would have published x studies, according to this inverse square law. In summary, the number of authors who produce *x* studies is proportional to $1/x^2$ .In a given field, 100 authors have published a single study. One can predict how many authors would have published twice, according to the following formula $y_x : c \times 1/x^2$ In this formula y_x is the number of authors with *x* publications, c is the number of authors with a single publication and x is the number of publications itself. Therefore, to find the number of authors who have contributed two articles (x =2) it is $y_2 = 100/2^2 = 25$. Thus it is predicted that 25 authors would have published two articles and so on. The number of authors will decrease in a progressive way, so only 11 authors would have published three articles according to formula ($y_3 = 11$). As the number of publications increases, the number of authors who have published χ articles decreases. Accordingly, 245 percent of contributions will correspond to 75 percent of the less productive authors, whereas the most productive authors will account for 50 percent of contributions. The present study focus on the authorship pattern and test lotka's law and Kolmogorov-Smimov (K-S) goodness-of-fit test are applied on author distribution data collected from the published literature of library and information science of the three chosen countries India, US and UK during the studied period 2007 to 2017.

Hypotheses Formulated

For the purpose of the study following null hypotheses have been formulated and tested based on data collected:

- *H_{o1}= There is a similar pattern of degree of collaboration among three different countries. That is, there is no association between authorship types and countries.*
- *$H_{o2:}$ The observed authorship data distribution of Indian literature is same as the theoretical authorship data distribution i.e. Lotka's law.*
- *$H_{o3:}$ The observed authorship data distribution of US literature is same as the theoretical authorship data distribution i.e. Lotka's law.*
- *$H_{o4:}$ The observed authorship data distribution of UK is same as the theoretical authorship data distribution i.e. Lotka's law.*
- *$H_{o4:}$ The observed authorship data distribution of the compiled data set (India, US and UK) is same as the theoretical authorship data distribution i.e. Lotka's law.*

Literature Review

Testing of the validity of Lotka's law has been performed. It has been found that Lotka's law can be satisfactorily applied to the literature brought out by the library science scholarly output in the selected journal literature of the four countries. The validity of Lotka's law was performed by Tamilselvan and Sivakumar in a study conducted for analyzing bibliometric literature published by faculties of National Institute of Technology (NIITs) in India. It was found that Lotka's law was satisfactorily applied to the articles published by NIT faculties.

Senthilkumar and Ulaganathan retrieved data related to Astrophysics information in India from web of science and analyzed various scientometric indicators were applied to total of 12750 astrophysics research output in India as D value 0.325 was found greater than critical value 0.120..Ahmad and Khan attempted to explain the theoretical aspects of the Lotka law. They applied it on the subject disciplines of Agricultural science and economics. The results did not showed compliance of Lotka's law. Ahmad and Rahman examined authorship distribution in the field of Nutrition in Bangladesh. The articles published during 1972-2006 were included. Lotka's generalized and modified law was applied with full author's productivity. The Lotka's generalized square law was not found applicable to the nutrition literature. Though the law holds well when the high productive authors were excluded from the group. Osraeh and Mostafawi collected 19,150 articles in the field of computer science produced by 45,713 authors from web of science database during 1986-2009. The Lotka's law and Kolmogorov-Smimov (K-S) goodness-of-fit test were applied. The study suggested certain points to be taken into consideration while testing lotka's applicability on a

particular dataset i.e. Lotka's original inverse square law or modified method , subject area, period of time, measurement of authors, estimation and testing of a criterion for assessing goodness-of-fit. Govinlaradjou and John assessed the research performance of the journal Ecology by analyzing its publication output during a period of ten years (2003-2012) using bibliometric indicators. Lotka's inverse law using Pao's method was tested. The calculated value of D (0.056) was found less compared to the table value of 0.019, thus conforming author distribution to Lotka's law.

Kumar and SenthilKumar investigated authors scientific research output in the field of Astronomy and Astrophysics and tested applicability of Lotka's law in generalized as well as inverse square form. The results of the statistical tests proved that Lotka's law does not fit the contribution frequency of author's productivity. Tsai analyzed the subject heading of Data Mining during the period covered from 1989 to 2009 in the SSCI database. The research findings showed that the main document type is the research articles. The results of the K-S tests proved that author productivity distribution fits Lotka's law in the subject of data mining. Muthukrishnan and SenthilKumar examined oncology research productivity in India for the period 2005-15. The web of science database data collected showed that the total number of authors i.e. 21443 therefore confirmed to lotka's law with D value=0.006 at 0.01 level of significance, with K-S statistics 0.011. Dhoble and Kumar analyzed 3588 papers on mustard research collected from CAB direct during 2000 to 2013. The results of the findings showed partial similarity between observed number and expected number of authors. Only one author contributed the highest number (50) of articles. As the number of articles increased the number of authors decreased. Sharma and Chakravarty tested fitness of Lotka's law on LIS literature of central universities in North India. Data was collected by online questionnaire, websites of universities, e-mails and phone calls. To validate Lotka's law, the calculated value of N and C are 0.01 and 0.04 respectively. The obtained D value 0.06 was less than the critical value 0.30 resulted in fitted Lotka's law.

Scope

The present investigation analyses publication trends and application of lotka's law to author productivity among LIS professionals who contributed journal articles in the chosen periodicals. Journals selected for inclusion in this study are prominent research oriented journals in the subject

of LIS. Further from each country India, US and UK four prestigious journals are chosen to conduct the study.

The research focuses on select LIS journals that publish peer-reviewed research articles. The investigations started with the most authoritative publisher's site Emerald library science subject collection accessible through UGC-INFONET consortium. During the search multiple websites were also searched to select indexed journals. The access policy for all the journals was checked to check their availability. Other websites and online free directories and databases which were consulted are as follows:

- **Directory of Open Access Journals (Lund University Library)**
- **LISTA**
- **SCOPUS**
- **Indian Journals.com**
- **LISA**

Finally, only those journals were selected for the study which were publishing articles since 2007 and continued publications. The journals which were easily accessible through UGC-INFONET consortium and were properly indexed were selected. The title of the journals is as follows refined sample. Firstly, the journals must have publications in English language only and also must have included a fair count of research articles. Secondly, these journals should be indexed and abstracted by proper indexing and abstracting services. Thirdly, these journals should have commenced their publication on or before 2007. From each of the three sampled countries four journals from each country fulfilling these criteria's were selected. Finally, there were 12 journals in the sample meeting these criteria. The titles of the selected journals are the following:

India

1. Annals of Library and Information Studies
2. DESIDOC Journal of Library and Information Technology
3. SRELS Journal of Information Management
4. Information Studies

US

1. Library Trends
2. Portal: Library and the Academy
3. College and Research Libraries
4. Information Technology and Libraries

UK

1. The Electronic Library
2. Journal of Documentation
3. Journal of Information Literacy
4. Library Management

All the above journals except Information Studies and SRELS Journal of Information Management have been accessible freely during the period covered 2007-2017;only the Indian journal " Information Studies " ceased its publication from 2015. So, it has been covered from 2007-2015 only. The standard abbreviated forms of names of the journals have been used. The standard terminology according to International Standard ISO 4 has been used. The mnemonic abbreviations were assigned for the sake of representing tabular data in limited space in the subsequent chapters. consists of general analysis of data related to twelve journals four each from the three countries India, US and UK representing data related to total number of authors, authorship pattern and collaboration trends along with collaboration coefficient of the sampled data collected. For testing of various hypotheses chi square test has been applied.

Table 1
Journal wise Author distribution

India			**US**			**UK**		
Journal name	**No. of Author**	**%**	**Journal name**	**No. of Author**	**%**	**Journal name**	**No. of Author**	**%**
Ann.Libr.Inf.Stud.	705	22.29	Coll Res. Libr	830	34.3	Electronic Lib.	1361	38.38
DESIDOCJ. Lib. Inf. Technol.	1062	33.59	Inf.Technol. Libr	309	12.77	J.Doc	1016	28.65
Inf. Stud.	281	8.89	Libr Trends.	728	30.08	J. Inf. Lit.	256	7.22
SRELS J.Inf. Manage.	1114	35.23	Portal	553	22.85	Libr.Manage.	913	25.75
Total	**3162**	**100**	**Total**	**2420**	**100**	**Total**	**3546**	**100**

Table 1 shows data related to total number of authors in the 12 journals and break up of author count in each of individual journal published from India, US and UK. In India with 3162 total authors , SRELS J. Inf. Manage., with 1114(35%) has highest number, while Coll Res. Libr., in the country US with 830(34.3%) count of authors has the largest number and Electron. Libr., UK with 1361(38.8%) has total count of authors.

Table 2
Country-wise authorship pattern

Authorship Type	India	US	UK
Single	601(35.49%)	660(50.26%)	757(41.61%)
Two	798(47.13%)	391(29.77%)	651(35.78%)
Collaborative	294(17.36%)	262(19.95%)	411(22.59)
Total	1693	1313	1819
Degree of collaboration	0.17	0.2	0.23

Table 2 represents single, joint and collaborative authorship pattern in the journal publications of India, US and UK. The publication literature in India shows trends towards collaborative authorship with two (47.13%) and joint (17.36%) collaborative authorship but only (35.49%) single authorship pattern. The US journal literature indicates maximum percent of single authorship papers i.e. (50.26%) followed by (29.77%) joint authors and (19.95%) more than two or collaborative papers. In UK scholarly literature (35.78%) are two author publications and (22.59%) are collaborative works while single author publications are (41.16%).

H_{o5}= There is a similar pattern of degree of collaboration among three different countries. That is, there is no association between authorship types and countries.

Table 2a
Authorship type and Country Cross tabulation

Count						**Chi-Square Tests**
		Country			**Total**	**Value**
		India	**US**	**UK**		
Type of Authorship	Single	601	660	757	2018	Pearson Chi-Square=66.331
	Collaborated	1092	653	1062	2807	df=2
Total		**1693**	**1313**	**1819**	**4825**	**Asymp. Sig. (2-sided)= .000**

*Significant at 0.05 level

Table 2a shows the value of chi square test performed as 66.331 with degree of freedom 2 and p value is less than 0.05 so at 95% confidence interval we reject the null hypothesis and conclude that

authorship type varies according to country i.e. there is an association between type of authorship and country. Indian and UK accounted for more number of collaborative works while US has large number of solo article contributions.

Table 3
Country wise collaboration trends

No. of Authors	India	US	UK
1 Author	601(35.49%)	660(50.26%)	757(41.61%)
2 Author	798(47.13%)	391(29.77%)	651(35.78%)
3 Author	230(13.58%)	167(12.71%)	248(13.63%)
4 Author	48(2.83%)	50(3.8%)	113(6.21%)
5 Author	13(0.76%)	21(1.59%)	28(1.53%)
6 Author	3(0.17%)	12(0.91%)	12(0.65%)
7 Author	0	6(0.45%)	9(0.49%)
8 Author	0	3(0.22%)	0
9 Author	0	2(0.15%)	0
15 Author	0	0	1(0.05%)
16 Author	0	1(0.07%)	0
Total	**1693**	**1313**	**1819**

Table 3 depicts in the published journal literature of the three countries the single author publications are highest in US i.e. 50.26% and thereafter UK with 41.61% and at last India 35.49%. Two author publications are largest in India 47.13% and after this come UK with 35.78% and US 29.77%. Three author publications are in large quantity in scholarly literature of UK 13.63%. India 13.58% and US 12.71%. UK has maximum number of four author publications 6.21% and then comes US with 3.8% and with least number of papers in India 2.83%. Five and six author publications are nearly the same in all the three countries. There are no articles with seven authors in India but UK and US has 0.49% and 0.45% respectively. Eight with 0.22% and nine authors 0.15% are associated with only US periodicals. There is only one i.e. 0.05 fifteen author publication in UK only 1(0.07%) paper authored by sixteen authors in US.

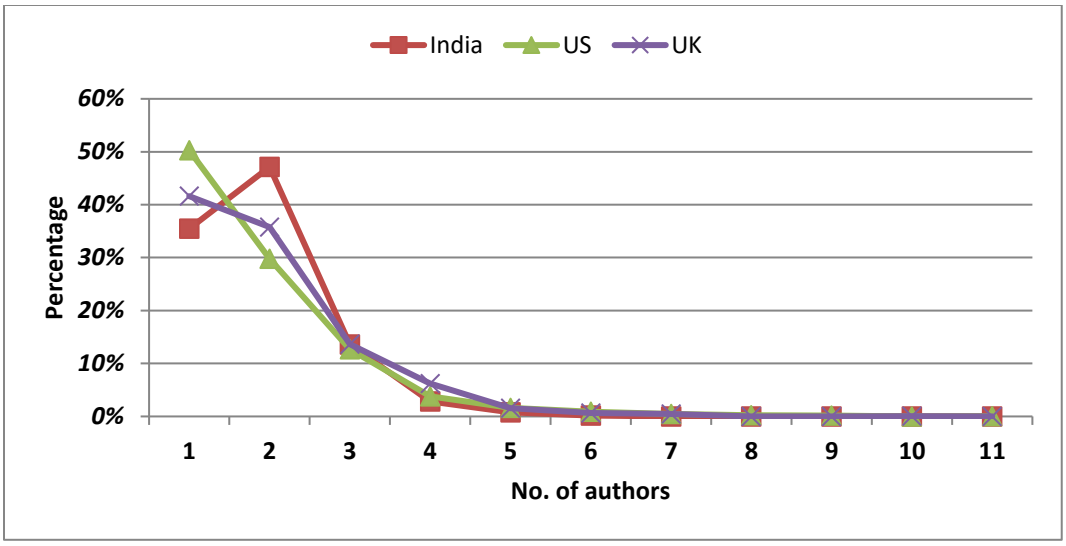

Figure 1.

Collaboration coefficient (CC)

The collaboration Coefficient (CC) has been calculated among the three countries India, US and UK suggested by Ajiferuke (1988). CC is a number between 0 and 1. If it is more than 0.5 it shows better collaboration among the authors. If it is near 0, it signifies weak collaboration. Collaborative Co-efficient: The collaborative coefficient has been calculated to observe the country-wise distribution of collaboration trends.

Country wise collaboration trends.

$$CC = 1 - \left[\sum_{j=1}^{k} \left(\frac{1}{j}\right) Fj/N\right]$$

Where,

Fj = the number of authored paper

N= Total number of research published; and

K=the greatest number of authors per paper.

India

$$=1-\left[\frac{\frac{1}{1}\times 601+\frac{1}{2}\times 798+\frac{1}{3}\times 230+\frac{1}{4}\times 48+\frac{1}{5}\times 13+\frac{1}{6}\times 3}{1693}\right]$$

$$=1-\left[\frac{601+399+76.66+12+2.6+0.5}{1693}\right]$$

$$=1-\frac{1091.76}{1693}=\frac{601.24}{1693}$$

CC=0.3551

US

$$=1-\left[\frac{\frac{1}{1}\times 560+\frac{1}{2}\times 391+\frac{1}{3}\times 167+\frac{1}{4}\times 50+\frac{1}{5}\times 21+\frac{1}{6}\times 12+\frac{1}{7}\times 6+\frac{1}{8}\times 3+\frac{1}{9}\times 2+\frac{1}{15}\times 0+\frac{1}{16}\times 1}{1313}\right]$$

$$=1-\left[\frac{560+195.5+55.66+12.5+4.2+2+0.85+0.375+0.22+0.625}{1313}\right]$$

$$=1-\frac{831.93}{1313}=\frac{481.07}{1313}$$

CC=0.3663

UK

$$=1-\left[\frac{\frac{1}{1}\times 757+\frac{1}{2}\times 651+\frac{1}{3}\times 248+\frac{1}{4}\times 113+\frac{1}{5}\times 28+\frac{1}{6}\times 12+\frac{1}{7}\times 9+\frac{1}{8}\times 0+\frac{1}{4}\times 0+\frac{1}{15}\times 1+\frac{1}{16}\times 0}{1819}\right]$$

$$1-\left[\frac{757+325.5+82.66+28.25+5.6+2+1.28+0.66}{1819}\right]$$

$$=1-\frac{1202.3}{1819}=\frac{1819-1202.3}{1819}=\frac{616.7}{1819}$$

CC=0.3390

The average collaboration co-efficient is 0.3535 in the journal literature of the three countries. This publication collaboration trend has been in tune with the study conducted in variety of disciplines like Economics Van Praag and Van Praag (2004) and Information science Al-Ghamdi et.al. (1998).

Lotka's Law

In 1926 Alfred J. Lotka published his research work representing frequency distribution of authorship pattern and scientific productivity observed from a decimal index of Chemical abstract during the period 1907-1916. He concluded that:-

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

Lotka's law serves as a pioneering work to test the author productivity in a particular field. It states that 60 percent of the authors will publish one article; 15 percent will have two article contributions ($1/2^2$ times 60); and only 7 percent will contribute 3 publications ($1/3^2$ times 60) and so on (Lotka,1926; Rowland,2005; Yueh et al. 2000).

The law of Lotka established that number of authors, Yx, each one of them 'x', is inversely proportional to x that is the productivity of each individual author.

The relationship is expressed as

$x^n . y=c$

Where y denotes the probability of an author to publish 'x' times in a particular field. n and c are the parameters to be estimated in a particular dataset.

The least square method proposed by (Pao, 1985) have been used to calculate the value of n

$$n=\frac{N\Sigma XY-\Sigma XY}{\Sigma X^2-(\Sigma X)^2} \quad \ldots\ldots \text{Formula (1)}$$

N= the number of pairs of data

X= the logarithmic value of articles (X)

Y= the logarithmic value of authors (Y)

The 'c' value which is the theoretical number of authors with a single article is calculated in the following way. The parameter C is calculated using the following formulae (Pao 1985)

$$C = \frac{1}{\left[\sum_{1}^{p=1} + \frac{1}{(n-1)(p^{n-1})} + \frac{1}{2p^{n}} + \frac{n}{24(p-1)^{n-1}}\right]} \quad \ldots\ldots\ldots\ldots\text{Formula (2)}$$

P=20, n is the value obtained using formulae (1); and x= number of articles.

Pao (1985) suggests that 'K-S test' a goodness of fit statistical test to assert that the observed author productivity distribution has not been significantly different from a theoretical distribution. To prove the hypotheses a comparison of the value is required. The test determines the associated probability that the observed maximum deviation occurs within limits of chance. The DMax that is maximum deviation between the cumulative proportions of the observed and theoretical frequencies is calculated using the following formulae (Pao, 1985):

$D = \text{Max}|F_o(\chi) - Sn(\chi)$

F0 (χ) =the theoretical cumulative frequency

Sn(χ) = the observed cumulative frequency

The test is performed at 0.01 level of significance, when sample size is greater than 35, the critical value of significance is calculated by following formulae. (Pao, 1985):

The critical value at the 0.01 level of significance:

$$C.V = \frac{1.63}{\sqrt{\Sigma yx + \sqrt{\frac{\Sigma yx}{10}}}} \quad \ldots\ldots\ldots \quad \text{Formula (3)}$$

If the maximum deviation falls within the critical value the null hypotheses that the data set conforms to Lotka's law can be accepted at a certain level of significance. But if it exceeds the critical value the null hypothesis must be rejected at a certain level of significance and concluded that the observed distribution is significantly different from the theoretical distribution.

Application of Lotka's law to analyze author productivity

In 1926, Lotka in his pioneering work carried out to test the regularity in the publication productivity concluded that:

" the number (of authors) making n contributions is about 1/n2 of those making one ; and the proportion of all contributors, that makes a single contribution, is about 60 percent".

The study analyses author productivity and examines the validity of Lotka's law on the individual author data sets of each country i.e. India, US and UK as well as compiled dataset of authors from all the three countries taken together. As this is a primary effort in just finding out the applicability of Lotka's law to a small piece of sample with conditions of considering only first authors, the efforts can be extended to multiple authors, authorship pattern considering other publication types, author affiliations etc.

$H_{o12a:}$ The observed authorship data distribution of Indian literature is same as the theoretical authorship data distribution i.e. Lotka's law.

Table 4
India: Author productivity and Kolmogorov-Smirnov test of observed and expected distribution

No. of Article	No. of Author	X=logx	Y=logy	XY	XX	y x / Σ y x	Σ (y x / Σ y x)	1/ x n	f = C(1/ x n)	Σ	D
1	963	0.0000	2.9836	0.0000	0.0000	0.2738	0.2738	1.0000	0.7103	0.7103	-0.4365
2	154	0.3010	2.1875	0.6585	0.0906	0.2007	0.4745	0.1993	0.1415	0.8518	-0.3773
3	58	0.4771	1.7634	0.8414	0.2276	0.1618	0.6364	0.0776	0.0551	0.9069	-0.2706
4	17	0.6021	1.2304	0.7408	0.3625	0.1129	0.7493	0.0397	0.0282	0.9351	-0.1859
5	9	0.6990	0.9542	0.6670	0.4886	0.0876	0.8368	0.0236	0.0168	0.9519	-0.1151
6	5	0.7782	0.6990	0.5439	0.6055	0.0641	0.9010	0.0155	0.0110	0.9629	-0.0619
7	4	0.8451	0.6021	0.5088	0.7142	0.0552	0.9562	0.0108	0.0077	0.9705	-0.0143
8	3	0.9031	0.4771	0.4309	0.8156	0.0438	1.0000	0.0079	0.0056	0.9762	0.0238
12	1	1.0792	0.0000	0.0000	1.1646	0.0000	1.0000	0.0060	0.0043	0.9804	0.0196
13	1	1.1139	0.0000	0.0000	1.2409	0.0000	1.0000	0.0047	0.0033	0.9838	0.0162
28	1	1.4472	0.0000	0.0000	2.0943	0.0000	1.0000	0.0038	0.0027	0.9865	0.0135
Total	**1216**	**8.2458**	**10.8974**	**4.3913**	**7.8043**	**1.0000**		**1.3888**	**0.9865**		

The above table 4 examines the validity of Lotka's law to author distribution on the subject of LIS, in the four periodicals published from India during 2007-2017. Lotka's law is applied and tested using Pao's method and verified through Kolmgorov-Smirnov (KS) test. Table 4 shows that almost(963) 79% produced single article,(154) 13% produced 2 articles, (58) 5% produced 3 articles,(17) 4% produced 4 articles, (9) 0.74% produced 5articles,(5) 0.41% produced 4 articles, (4) 0.32% gave

3 articles, (3) 0.24% produced 8articles, (1)0.08% authors produced 12 ,13 and even 28 total number of articles. It can be observed from the data that gradually as the number of articles are increasing number of authors are decreasing in the studied journal literature of India.

To validate Lotka's law calculations are done using the formulaes to identify whether values of '*n*' and 'C' to test whether application of Lotka's law fits into the data of present study or not.

From Table 4, shows the estimated value 'n' for the dataset is calculated using formulae (1).

$$n=\frac{N\Sigma XY-\Sigma XY}{\Sigma X^2-(\Sigma X)^2}$$

$$n=\frac{(11*4.391257)-(8.245803*10.89742)}{11*7.804349-(8.245803)^2} = -2.3273$$

The value of n of the Indian LIS research through Least Square Method produces a value of n= 2.3273

Calculation of constant 'c'

The calculated value of constant 'c' for LIS authors from Indian publications is0.713, obtained from the following formulae (2):

$$\sum \frac{1}{x^{2.3273}} = 1.40777$$

$$= \frac{1}{1.40777} = 0.7103$$

The constant c for the dataset is calculated using the formulae (3) and the value of c=0.7103. The critical value is calculated as 0.046.

$$C.V=\frac{1.63}{\sqrt{1216+\sqrt{\frac{1216}{10}}}} = 0.046$$

The Dmax value is the maximum difference between real and accumulated frequencies observed from the (Table 4) is 0.0238 and is lower than the critical value of Kolmogorov-Smirnov test at the 0.05 significance level i.e. 0.046 determined with n= 2.3273 and hence supports the consideration of the hypothesis that is the observed authorship data distribution holds good for the Lotka's law.

The authorship data distribution of LIS journal articles published in 4 Indian journals is being tested for the application of the Lotka's law. The hypothesis assumed that the observed data distribution is

same as the theoretical data distribution. The value n is determined through least square method. The calculated data is verified through Kolmogorov Smirnov test for various values of n. The observed distribution is also tested against the inverse square law using the exponent n=2.3273, it is found Indian journal literature do conform to Lotka's law.

H_{o2}: The observed authorship data distribution of US literature is same as the theoretical authorship data distribution i.e. Lotka's law.

Table 5
US: Author productivity and Kolmogorov-Smirnov test of observed and expected distribution

No. of Article	No. of Author	X=logx	Y=logy	XY	XX	y x / Σ y x	Σ (y x / Σ y x)	1/ x n	f = C(1/ x n)	Σ	D
1	1124	0.0000	3.0508	0.0000	0.0000	0.9328	0.9328	1.0000	0.9350	0.9350	-0.0022
2	61	0.3010	1.7853	0.5374	0.0906	0.0506	0.9834	0.0552	0.0516	0.9866	-0.0032
3	14	0.4771	1.1461	0.5468	0.2276	0.0116	0.9950	0.0101	0.0095	0.9961	-0.0010
4	5	0.6021	0.6990	0.4208	0.3625	0.0041	0.9992	0.0030	0.0028	0.9989	0.0003
5	1	0.6990	0.0000	0.0000	0.4886	0.0008	1.0000	0.0012	0.0011	1.0000	0.0000
Total	**1205**	**2.0792**	**6.6812**	**1.5051**	**1.1693**	**1.0000**		**1.0695**	**1.0000**		**0.0003**

Table 2 shows that almost (1124) 0.93%produced single article, (61) 5.06% produced 2 articles, (14) 1.16 % produced 3 articles, (5) 0.4% produced 4 articles, (1) 0.08% produced 5articles. It can be observed from the data that gradually as the number of articles are increasing number of authors are decreasing in the studied journal literature of US.

From Table 5 the estimated value 'n' for the dataset is calculated using formulae (1). The value of n of the US LIS researchers through Least Square Method produces a value of n= 4.18

The estimated value 'n' for the dataset is calculated as follows:

$$n=\frac{N\Sigma XY-\Sigma XY}{N\Sigma X^2-(\Sigma X)^2}$$

$$n=\frac{(5*1.505)-(2.0791*6.6812)}{5*1.1693-(2.0792)^2} = -4.145$$

The value of n through Least Square Method produces a value of n= 2.3273

Calculation of constant 'c'

The calculated value of constant ‘c’ for authors taken from US periodicals is 1.069, obtained from the following formulae:

$$\sum \frac{1}{x^{4.145}} = 1.069$$

$$= \frac{1}{1.06902} = 0.935$$

The constant c for the dataset is calculated using the formulae (2) and the value of c=0.935. The calculated critical value is 0.044.

$$C.V = \frac{1.63}{\sqrt{1205 + \sqrt{\frac{1205}{10}}}} = 0.044$$

From Table 5, the estimated value ‘n’ for the US dataset is calculated using formulae (1). The value of n of the LIS researchers through Least Square Method produces a value of n= 0.17

The Dmax value is 0.1362, which is the the maximum difference between the real and accumulated frequencies (Table 5) and is high than the critical value of Kolmogorov-Smirnov test at the 0.05 significance level i.e. 0.04 determined with n= 4.18 and hence does not supports the consideration of the hypothesis that is the observed US authorship data distribution holds good for the Lotka’s law.

H_{o3}: The observed authorship data distribution of UK is same as the theoretical authorship data distribution i.e. Lotka’s law.

Table 6
UK:Author productivity and Kolmogorov-Smirnov test of observed and expected distribution

No. of Article	No. of Author	X=logx	Y=logy	XY	XX	y x / Σ y x	Σ (y x / Σ y x)	1/ x n	f = C(1/ x n)	Σ	D
1	1309	0.0000	3.1169	0.0000	0.0000	0.8640	0.8640	1.0000	0.6120	0.6120	0.2520
2	148	0.3010	2.1703	0.6533	0.0906	0.0977	0.9617	0.2606	0.1595	0.7715	0.1902
3	39	0.4771	1.5911	0.7591	0.2276	0.0257	0.9874	0.1187	0.0726	0.8441	0.1433
4	10	0.6021	1.0000	0.6021	0.3625	0.0066	0.9940	0.0679	0.0416	0.8857	0.1083
5	4	0.6990	0.6021	0.4208	0.4886	0.0026	0.9967	0.0441	0.0270	0.9127	0.0840
6	2	0.7782	0.3010	0.2342	0.6055	0.0013	0.9980	0.0309	0.0189	0.9316	0.0664
7	1	0.8451	0.0000	0.0000	0.7142	0.0007	0.9987	0.0229	0.0140	0.9456	0.0530

8	1	0.9031	0.0000	0.0000	0.8156	0.0007	0.9993	0.0177	0.0108	0.9565	0.0429
10	1	1.0000	0.0000	0.0000	1.0000	0.0007	1.0000	0.0141	0.0086	0.9651	0.0349
Total	**1515**	**5.605521**	**8.781356**	**2.669573**	**4.304581**	**1.0000**		**1.5769**	**0.9651**		

Table 6 shows that almost(1309) 86% produced single article,(148) 9.76% produced 2 articles, (39) 2.57% produced 3 articles,(10) 0.66% produced 4 articles, (4) 0.26% produced 5articles,(2) 0.13% produced 6 articles, (1) 0.06% gave 7 articles, (3) 0.06% produced 8articles, (1)0.06% authors produced 10 total number of articles. It can be observed from the data that gradually as the number of articles are increasing number of authors are decreasing in the studied journal literature of India.

From Table 6, the estimated value 'n' for the dataset is calculated using formulae (1). The value of n of the UK LIS research through Least Square Method produces a value of n= 1.94

$$n=\frac{N\Sigma XY-\Sigma XY}{\Sigma X^2-(\Sigma X)^2}$$

$$n=\frac{(10*2.67)-(5.6055*8.7814)}{10*4.304-(5.6055)^2} = -1.93803$$

The value of n of the LIS researchers from UK periodicals through Least Square Method produces a value of n= 1.93803

Calculation of constant 'c'

The calculated value of constant 'c' for LIS is 0.612, obtained from the following formulae:

$$\sum \frac{1}{x^{1.93803}} = 1.634154$$

$$= \frac{1}{1.634154} = 0.612$$

The constant c for the dataset is calculated using the formulae (2) and the value of c=0.612. The calculated critical value is 0.046.

$$C.V=\frac{1.63}{\sqrt{1515+\sqrt{\frac{1515}{10}}}} = 0.046$$

The Dmax value is the difference calculated between the real and accumulated frequencies from the (Table 6) is 0.2520 and is higher than the critical value of Kolmogorov-Smirnov test at the 0.05 significance level is 0.046 determined with n= 1.94 and hence does not supports the consideration of the hypothesis that is the observed authorship data distribution holds good for the Lotka's law.

H_{o4}: The observed authorship data distribution of overall dataset of the three countries (India, US and UK) is same as the theoretical authorship data distribution i.e. Lotka's law.

Table 7
Compiled dataset of India, US and UK: Author productivity and Kolmogorov-Smirnov test of observed and expected distribution

No. of Article	No. of Author	X=log x	Y=log y	XY	XX	y x / Σ y x	Σ (y x / Σ y x)	1/ x n	f = C(1/ x n)	Σ	D
1	3396	0.0000	3.5310	0.0000	0.0000	0.8628	0.8628	1.0000	0.1100	0.1100	0.7528
2	363	0.3010	2.5599	0.7706	0.0906	0.0922	0.9550	0.8888	0.0978	0.2078	0.7473
3	111	0.4771	2.0453	0.9759	0.2276	0.0282	0.9832	0.8296	0.0913	0.2990	0.6842
4	32	0.6021	1.5052	0.9062	0.3625	0.0081	0.9914	0.7900	0.0869	0.3859	0.6054
5	14	0.6990	1.1461	0.8011	0.4886	0.0036	0.9949	0.7606	0.0837	0.4696	0.5253
6	7	0.7782	0.8451	0.6576	0.6055	0.0018	0.9967	0.7374	0.0811	0.5507	0.4460
7	5	0.8451	0.6990	0.5907	0.7142	0.0013	0.9980	0.7183	0.0790	0.6297	0.3682
8	4	0.9031	0.6021	0.5437	0.8156	0.0010	0.9990	0.7022	0.0772	0.7070	0.2920
10	1	1.0000	0.0000	0.0000	1.0000	0.0003	0.9992	0.6883	0.0757	0.7827	0.2165
12	1	1.0792	0.0000	0.0000	1.1646	0.0003	0.9995	0.6761	0.0744	0.8571	0.1424
13	1	1.1139	0.0000	0.0000	1.2409	0.0003	0.9997	0.6652	0.0732	0.9302	0.0695
28	1	1.4472	0.0000	0.0000	2.0943	0.0003	1.0000	0.6652	0.0732	1.0034	-0.0034
Total	**3936**	**9.2458**	**12.9336**	**5.2458**	**8.8043**	**1.0000**		**9.1220**	**1.0034**		**0.7528**

Table 7 describes the author productivity and Kolmogorov-Smirnov test of observed and expected distribution. All the authors are considered from the journal literature of the three countries India, US and UK. Table 4 shows that almost(3396) 86% produced single article,(363)10.6% produced 2 articles, (111)2.82 % produced 3 articles,(32)0.81% produced 4 articles, (14) 0.35% produced 5 articles,(7)0.17% produced 6 articles, (5)0.12% produced 7 articles, (4)0.1% produced 8 articles, (1)0.02% produced 10,12 and 28 total number of articles. It can be observed from the data that gradually as the number of articles are increasing number of authors are decreasing in the studied journal literature of the compiled author data set.

From Table 7, the estimated value 'n' for the dataset is calculated using formulae (1). The value of n of the LIS researchers through Least Square Method produces a value of n= 0.17

The estimated value ‘n’ for the compiled dataset is calculated as follows:

$$n=\frac{N\Sigma XY-\Sigma XY}{\Sigma X^2-(\Sigma X)^2}$$

$$n=\frac{(28*5.246)-(9.245*12.934)}{28*8.804-(9.246)^2} = 0.169$$

Calculation of constant ‘c’

The calculated value of constant ‘c’ for LIS is 0.7175, obtained from the following formulae:

$$\sum \frac{1}{x^{0.169}} = 1.3937$$

$$= \frac{1}{1.3937} = 0.7175$$

The constant c for the dataset is calculated using the formulae (2) and the value of c=0.7175. The calculated critical value is 0.03

$$C.V=\frac{1.63}{\sqrt{3936+\sqrt{\frac{3936}{10}}}} = 0.03$$

The Dmax value is 0.7528, the maximum difference calculated between the real and estimated accumulated frequencies. (Table 7) and is higher than the critical value of Kolmogorov-Smirnov test at the 0.05 significance level is 0.03 determined with n= 0.17 and hence does not supports the consideration of the hypothesis that is the observed authorship data distribution holds good for the Lotka’s law.

Conclusion

- Total number of authors in the 12 journals and author count in each of the individual journal published from India, US and UK has shown that in India 3162 total authors , SRELS J. Inf. Manage., with 1114(35%) has highest number, while Coll Res. Libr., in the country US with 830(34.3%) has the largest count and Electron. Libr., from UK with 1361(38.8%) has the largest share of authors .
- A trend in academia shows rise in the level of publication collaboration as proved in variety of academic fields like economics (Van Praag and Van Praag 2004) and information science (Al-Ghamdi et al. 1998). Country-wise single and collaborative authorship pattern. The

Indian periodical literature studied shows inclination of authors towards two 47.13% and joint authorship 17.36%. While single authors are 35.49% only. The UK literature shows the same trend with good number of single authored articles 41.61% but when two author articles i.e. 35.78% and 22.59% collaborative papers are taken together they outnumber solo author publications. The US research publications demonstrate 50.26% single author publications followed by two 29.77% and 19.55% collaborative works . The results of the sample coincides with Al-Ghamdi et al. (1998) who found increased participation of women as authors with increased collaboration in information science. The data analyzed gives the preference to collaborative effort in India and UK. That is why degree of collaboration is high in UK (0.23) and in India (0.17) but less in US (0.2) only. It is proved by the value of chi square is 66.331 with degree of freedom 2 and p value is less than 0.05 so at 95% confidence interval that authorship type varies according to countries i.e. there is an association between type of authorship and countries.

- In the published journal literature of the three countries the single author publications are highest in US i.e. (50.26%) and thereafter UK with (41.61%) and at last India (35.49%).Two author publications are largest in India (47.13%) and after this come UK (35.78%) and US (29.77%). Three author publications are in large quantity in scholarly literature of UK (13.63%), India (13.58%) and US (12.71%). UK has maximum number of four author publication (6.21%) and then US with (3.8%) and with least number of papers in India (2.83%).Five and six author publications is nearly the same in all the three countries. There are no articles with seven authors in India but UK and US has (0.49%) and (0.45%) respectively. Eight (0.22%) and nine authors (0.15%) are associated with only US periodicals. There is only one (0.05) fifteen author publication in US and 1(0.07%) paper authored by sixteen authors.
- The study analyses author productivity and examines the validity of Lotka's law on the individual author data sets of each country i.e. India, US and UK as well as on compiled dataset of authors from all the three countries taken together. The hypothesis assumed that the observed data distribution is same as the theoretical data distribution. The value n is determined through least square method. The calculated data is verified through Kolmogorov Smirnov test for various values of n.

- The observed authorship data distribution of Indian literature is same as the theoretical authorship data distribution i.e. Lotka's law. The Dmax value is the maximum difference between real and accumulated frequencies observed from the (Table 4) is 0.0238 and is lower than the critical value of Kolmogorov-Smirnov test at the 0.05 significance level i.e. 0.046 determined with n= 2.3273 and hence supports the consideration of the hypothesis that is the observed authorship data distribution holds good for the Lotka's law.
- The authors' productivity data set in US periodicals with Dmax value 0.1362 is found greater than the critical value of K-S test is 0.04 determined with n= 4.18 , hence hypothesis (Ho_3) is rejected as it does not supports the consideration of that is the observed authorship data distribution holds good for the Lotka's law .
- The formulated hypothesis that observed authorship data distribution of UK is same as the theoretical authorship data distribution i.e. Lotka's law. The Dmax value 0.2520 exceeds the critical value of K-S test i.e. 0.04 determined with n= 1.94 and hence the hypothesis is rejected. The observed authorship data distribution of overall dataset of the three countries (India, US and UK) with Dmax value 0.7528 proved to be more than the critical value of i.e. 0.03 determined with n= 0.17 and hence does not supports the consideration of the hypothesis that is the observed authorship data distribution holds good for the Lotka's law.

References

Ahmad, S., Ullah, S., Khan, F., & Lotka, A. J. (2012). Lotka ' s Law of Scientific Productivity : Theory and Application. *PUTAJ Humanities and Social Sciences*, 127–135.

Ajiferuke, I., Burell, Q., & Tague, J. (1988). Collaborative coefficient: A single measure of the degree of collaboration in research. *Scientometrics*, *14*(5–6), 421–433. https://doi.org/10.1007/BF02017100

Ali Al-Ghamdi, Mohsin Al-Harbi, Natalie A. B. Beacom, Jennifer Dedolph, Marci Diegnan, Charles lftmann, Nancy Finley, Leann Locicero, John Middlecamp, Christine O?Regan, Faymarie Pluskota, Andrew A.R itter Scott Russell, I., & Sabat, Joseph Schneider, Maria Schoeberl, Phyllis Tragash, and B. H. Wi. (1998). Authorship in JASIS: A quantitative analysis. *Katharine Sharp Review*, *Winter*(6), 190–206.

Dhoble, S. (2017). Applicability of Lotka ' s law in Mustard research publications in India - a scientometric study. *SSARSC International Journal of Library, Information Networks and Knowledge*, *2*(1), 1–10.

Govindaradjou, S., & John, D. (2014). Quantitative analysis of research trends in a leading

ecological journal: bibliometric study during 2003-2012. *South African Journal of Libraries and Information Science, 80*(1). https://doi.org/10.7553/80-1-178

Goyena, R., & Fallis, A. . (2019). Summary for Policymakers. In Intergovernmental Panel on Climate Change (Ed.), *Climate Change 2013 - The Physical Science Basis* (Vol. 53, pp. 1–30). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107415324.004

Kumar, S., & Senthilkumar, R. (2019). Applicability of Lotka's law in astronomy & astrophysics research of India. *Library Philosophy and Practice*, *2019*(January).

Kumaresan, R., Ezhilrani, R., & Vinitha, K. (2014). Lotka's Law: A study with reference to the literature by faculities of National Institutes of Technology, India. *International Journal of Library and Information Science (IJLIS)*, *2*(1), 19–34.

Lotka, A.J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences, 16,* 317-323.

Muthukrishnan, M., & Senthilkumar, R. (2017). Author productivity of oncology research output in India: Testing lotka's law. *International Journal of Information Dissemination and Technology, 7*(3), 187. https://doi.org/10.5958/2249-5576.2017.00021.8

Osareh, F., & Mostafavi, E. (2011). Lotka's Law and authorship distribution in Computer Science using Web of Science (WoS) during 1986–2009. *Collnet Journal of Scientometrics and Information Management, 5*(2), 171–183. https://doi.org/10.1080/09737766.2011.10700911

Pao, M.L. (1985). Lotka's law: A testing procedure. Information Processing & Management, 21(4), 305-320

Senthilkumar R, U. G. (2016). Scientometric analysis of Astrophysics research output in India. *Kongunadu Research Journal*, *3*(1), 84–87. https://doi.org/10.26524/krj136

Tsai, H.-H. (2012). Global data mining: An empirical study of current trends, future forecasts and technology diffusions. *Expert Systems with Applications*, *39*(9), 8172–8181. https://doi.org/10.1016/j.eswa.2012.01.150

Van Praag, M. C., & Van Praag, B. M. S. (2008). The benefits of being economics professor A(rather than Z). *Economica*, *75*(300), 782–796. https://doi.org/10.1111/j.1468-0335.2007.00653.x

Zabed Ahmed, S., & Anisur Rahman, M. (2009). Lotka's law and authorship distribution in nutrition research in Bangladesh. *Annals of Library and Information Studies, 56*(2), 95–102.