Application of Statistical Models to the Collaborative Publications in Corona Virus

R. Natarajan
Assistant professor
Dept. of Library &Information Science
DDE Wing, Annamalai University,
Annamalnagar-608002

M. Umamaheswari
Research Scholar
Dept. of Library &Information Science
DDE Wing, Annamalai University,
Annamalnagar-608002

Abstract

This article deals with the growth and proportion of different types of co-authored publications in Corona Virus. This study also explores the applicability of appropriate statistical model to the deterioration in the proportion of single-authored publication with during sample periods. Also studies the applicability of selected statistical models to the distribution of authorship in publications of corona virus during 1989 to 2014

Keywords

Collaboration; Publications; Statistical models; Corona virus, authorship pattern
.

Electronic access

The journal is available at www.jalis.in

Journal of Advances in Library and Information Science
ISSN: 2277-2219 Vol. 4. No.2. 2015. pp. 172-176

INTRODUCTION

Global wise majority of science and social science research has come through collaborative pattern. According to Beaver and Rosen (1978), The first collaborated papers has published in 1655, since then the proportion of total collaborated papers has been increasing with time. It was during the 20th century that the professionalization in science had its maximum impact on the members of the scientific community. Because of this factor, there had been an increasing trend in the growth of collaborative publications in all disciplines of science and technology (S&T). However, the increase in the number and proportion of collaborative publications and their growth rates have been observed to vary from one subject to another, one branch to another, within the same subject, and from one country to another. Changes in the growth rate of collaborative publications observed in subject disciplines are probably related to changes in the acknowledgement of teamwork, as an accepted practice in collaboration that simultaneously effect changes in the support of S&T.

OBJECTIVES

The main objectives of this study are:

(a) To study the growth in the number and proportion of collaborative publications in the field of Corona virus research with time;
(b) To study the applicability of a suitable mathematical model in the decline of the proportion of single-authored publications with time; and
(c) To study the applicability of selected probability distributions to the distribution of authorship with time for publications in the field of Corona virus.

DATABASE AND METHODOLOGY

The present study covers publication from 1989 to 2014 (26 years). For studying the growth in the number and proportion of multi-authored publications in the field of corona virus, the entire data from the bibliography was used. Corona virus research articles were downloaded from the database of Web of Science during the period 1989 to 2014, it includes the databases of SCI (Science Citation Index), SSCI (Social Science Citation Index) and AHCI (Arts and Humanities Citation Index)). The search key term has ‘corona virus’ has been used for

the purpose of collection of data, required for the study. It covers different types of publications in the discipline of corona virus during the sample periods. Further, the data was tabulated using by the software's of Histcite, and MS-Excel and further analyzed for the purpose of interpretation and discussion.

DISCUSSION AND RESULTS

Overall Collaboration Profile in Theoretical Population Genetics

The growth in the number of total publications and collaborative publications have considered shown a consistent increasing trend with time since 1989. However, the proportion of collaborative publications in total publications has shown a reliable growth with slight fluctuations in certain periods. The proportion of single-authored publications has increased from 0.35 percent during 1989 to 5.39 percent during 2014. Single authored contribution has shown in declining trend. The proportion of multi-authored publications has increased from 0.87 percent during 1989 to 8.11 percent during 2014. Multi-authored contribution shows in continuous increasing trend.

Some mathematical measures have been proposed by scholars in the past to study the extent and size of co-authorship, as reflected in publications. These measures are Degree of Collaboration (DC) first suggested by Subramanyam (1983), which takes the proportion of co-authored publications in total publications, Collaboration Index (CI), which takes the mean number of authors per publication, and Degree of collaboration first suggested by Ajiferuke, Burrell, and Tague (1988) that takes the proportional mean of the sum of publications and number of authors, and set the values between 0 and 1. In order to study the extent of collaboration in different period in Corona virus, the values of DC, and CI were computed for publications and the results obtained are presented in Table 1.

Table 1: Total Publications and Collaborative Publications between 1989 to 2014

Year	Single authors		Multi authored		Total	Degrees of Collaboration	Collaborative index
	No of output	%	No. of output	%			
1989	2	0.35	74	0.87	76	0.97	3.54
1990	10	1.74	85	1.0	95	0.89	3.59
1991	8	1.39	149	1.76	157	0.95	3.81
1992	6	1.04	143	1.69	149	0.96	3.98
1993	19	3.30	136	1.61	155	0.88	3.59
1994	9	1.57	139	1.64	148	0.94	3.88
1995	16	2.78	159	1.88	175	0.91	3.87
1996	12	2.09	136	1.61	148	0.92	3.82
1997	9	1.57	159	1.88	168	0.95	3.76
1998	8	1.39	190	2.25	198	0.96	4.20
1999	17	2.96	141	1.67	158	0.89	4.06
2000	8	1.39	137	1.62	145	0.94	4.23
2001	15	2.61	200	2.36	215	0.93	4.09
2002	5	0.87	139	1.64	144	0.97	4.56
2003	47	8.17	338	3.99	385	0.88	6.41
2004	58	10.09	725	8.57	783	0.93	7.14
2005	40	6.96	681	8.05	721	0.94	7.39
2006	33	5.74	686	8.11	719	0.95	6.18
2007	38	6.61	525	6.20	563	0.93	6.24
2008	30	5.22	519	6.13	549	0.95	5.95
2009	16	2.78	460	5.44	476	0.97	5.97
2010	22	3.83	450	5.32	472	0.95	6.07
2011	19	3.30	401	4.74	420	0.95	6.47
2012	24	4.17	447	5.28	471	0.95	6.43
2013	73	12.7	557	6.58	630	0.88	6.51

2014	31	5.39	686	8.11	717	0.96	6.98
Total	**575 (6.36)**	**100**	**8462 (93.64)**	**100**	**9037**	**0.94**	**5.93**

The year wise distribution of research analysis, the year of 2004 has highest number of publications 783 (8.7 %) with 15197 (13.55 %) of TLCS and 26280 (11.81 %) of TGCS values were scaled and being a first position among the 26 years output in the Corona virus research at comprehensive level. The year of 2005 has 721 (8.0 %) records and it stood in second position with 12772 (11.39 %) of TLCS and 23678 (10.64 %) of TGCS were scaled. Followed by 2006 has 719 (8.0 %) of publications in third position along with 7421 (6.62 %) of TLCS and 17852 (8.02 %) of TGCS were measured. Next to that 2014 has 717 (7.9 %) records with 571 (0.51 %) of TLCS and 1498 (0.67 %) of TGCS; and the remaining years were produced below 700 number of articles in corona virus research.

The overall periods Corona virus productivity has 6.36 percent of articles produced by the single authors and 93.64 percent of articles were produced by the collaborative authors. The collaborative value is 0.97 at 1989 and its raised 0.96 at 2014. The degree of collaboration value is 0.94, i.e. 94 percentage articles were produced by the collaborative work. The collaborative index has calculated based on the number of contributed authors. The following table analysis result has given elaborately the collaborative index value. The Collaborative Index values, it can be obtained by the total number of authors divided by the total number of published articles.

The following results were; such that: the year of 1989 has CI value is 3.54; year of 1990 has CI value is 3.59; 1991 has CI value is 3.81; year of 1992 has CI value is 3.98; year of 1993 has CI value is 3.59; year of 1994 has CI value is 3.88; year of 1995 has CI value is 3.87; year of 1996 has CI value is 3.82; year of 1997 has CI value is 3.76; year of 1998 has CI value is 4.20; year of 1999 has CI value is 4.06; year of 2000 has CI value is 4.23; year of 2001 has CI value is 4.09; year of 2002 has CI value is 4.56; year of 2003 has CI value is 6.41; year of 2004 has CI value is 7.14; year of 2005 has CI value is 6.18; year of 2006 has CI value is 6.24; year of 2007 has CI value is 6.24; year of 2008 has CI value is 5.95; year of 2009 has CI value is 5.97; year of 2010 has CI value is 6.07; year of 2011 has CI value is 6.47; year of 2012 has CI value is 6.43; year of 2013 has CI value is 6.51 and 2014 has CI value is 6.98 respectively. The collaborative Index for universal level value is 5.93 which show the status towards collaborative research pattern than single research in Corona virus.

Quantification of Publications by Number of Authors

In the total publication sample, Totally 9037 articles were produced by different types authorship pattern during 1989 to 2014; among those 575 (6.36 %) of articles by single authored, 1044 (11.55 %) of articles by two authored team, 1203 (13.31 %) of articles by three authored team, 1159 (12.83 %) of articles by four authored team; 1083 (11.98 %) of articles by five authored team, 892 (9.87 %) of articles by six authored team, 715 (7.91 %) of articles by seven authored team, 568 (6.29 %) of articles by eight authored team, 434 (4.80 %) of articles by nine authored team, 1353 (14.97 %) of articles by ten and above authored team. It could be identified the three authored production is highest than other team. Analyzing the percentage contribution of individual types of co-authored publications, a systematic increase is observed in all the categories of co-authored publications, although with different growth rates as described below:

Table 2: Publications classified by number of authors in Corona virus research during 1989 – 2014

Articles/Year	**1 Author**	**2 Authors**	**3 Authors**	**4 Authors**	**5 Authors**	**6 Authors**	**7 Authors**	**8 Authors**	**9 Authors**	**10 & above**	**TOTAL**
1989	2	18	23	14	11	5	3	0	0	0	76
1990	10	18	26	22	8	3	4	-	1	3	95
1991	8	29	35	28	19	13	6	4	2	1	157
1992	6	29	33	30	23	13	9	3	1	2	149
1993	19	29	34	34	18	8	6	5	1	1	155
1994	9	29	34	35	14	13	6	5	0	3	148

1995	16	30	38	39	21	12	9	3	4	3	175
1996	12	24	33	38	19	8	8	3	1	2	148
1997	9	31	40	36	30	10	10	1	1	0	168
1998	8	36	40	35	28	24	14	7	6	0	198
1999	17	28	31	22	22	14	9	8	5	2	158
2000	8	22	28	25	25	20	9	5	1	2	145
2001	15	41	45	32	34	20	10	10	4	4	215
2002	5	24	23	24	25	21	10	3	3	6	144
2003	47	56	47	40	41	30	15	16	15	78	385
2004	58	70	74	85	76	63	60	56	41	200	783
2005	40	66	54	75	67	83	63	47	46	180	721
2006	33	70	93	81	105	64	60	54	36	123	719
2007	38	43	63	65	63	63	56	48	34	90	563
2008	30	49	66	78	69	46	55	39	29	88	549
2009	16	50	51	57	60	62	47	43	31	59	476
2010	22	48	52	48	47	67	51	41	29	67	472
2011	19	40	41	39	60	46	38	35	29	73	420
2012	24	34	49	50	57	55	45	50	27	81	471
2013	73	57	72	60	60	57	46	38	39	128	630
2014	31	73	78	67	81	72	66	44	48	157	717
Total articles	**575 (6.36)**	**1044 (11.55)**	**1203 (13.31)**	**1159 (12.8)**	**1083 (11.98)**	**892 (9.87)**	**715 (7.91)**	**568 (6.29)**	**434 (4.80)**	**1353 (14.97)**	**9037**
Authors	**575 (1.07)**	**2088 (3.9)**	**3609 (6.74)**	**4636 (8.65)**	**5415 (10.11)**	**5352 (9.99)**	**5005 (9.34)**	**4544 (8.48)**	**3906 (7.29)**	**18445 (34.43)**	**53575**

Totally 53575 authors were contributed to the 9037 sample data during 1989 to 2014; among those 575 (1.07 %) were single authors; 2088 (3.90 %) were two authors, 3609 (6.74 %) were three authors, 4636 (8.65 %) were four authors, 5415 (10.11 %) were five authors, 5352 (9.99 %) were six authors, 5005 (9.34 %) were seven authors, 4544 (8.48 %) were eight authors, 3906 (7.29 %) were nine authors, and 18445 (34.43 %) were more than ten authors respectively. It could be identified the five authors, ten and above authors contribution is highest than other team. Thus, this analysis indicating that very clearly the increased trend towards on multi-authored productivity in the field of corona virus.

Prolific authors

Out 53575 authors, "Yuen KY**"** has published 152 articles contributed highest number of publications in Corona virus research with 7127 local citation scores; 10195 global citation scores; 3524 Total local citation reference values were measured and it occupies the first rank in Corona virus research output. Next to that "Enjuanes L**"** has published 137 articles with 2887 local citation scores; 4154 global citation scores; 3335 Total local citation reference values measured; "Perlman S**"** has published 116 articles with 1961 LCS; 3051 GCS; 2017 TLCR with values were measured. Remaing authors were contributed below 70 articles.

Table 3: Showing the prolific authors in Corona virus research output

S.No	Author name	Recs.	Rank	TLCS	TGCS	TLCR
1	Yuen KY	152	1	7127	10195	3524
2	Enjuanes L	137	2	2887	4154	3335
3	Perlman S	116	3	1961	3051	2017
4	Rottier PJM	113	4	2906	4908	2348
5	Weiss SR	111	5	2339	3267	2614
6	Baric RS	103	6	2054	2967	2501

7	Chan KH	100	7	5596	7859	1786
8	Drosten C	95	8	3962	5398	1534
9	Holmes KV	93	9	2757	4233	1545
10	Snijder EJ	92	10	2620	5491	2177

Table 4: Showing Bradford's Distribution of sources in Corona virus

Zone	No. of sources	No. of records	Multiplier factor
Z_1	15 (1.22)	3055 (33.80)	-
Z_2	101 (8.24)	2968 (32.84)	6.73
Z_3	1109 (90.53)	3014 (33.35)	10.98
Total	**1225**	**9037**	**17.71 (8.85)**

The researcher has grouped three types, such that; Z_1 – Zonal 1, Z_2 – Zonal 2 and Z_3 – Zonal 3. The first zone indicates the small groups of different types of sources, and it is called the nuclear or core zone. Here from this analysis, first zone representing 15 (1.22 %) of sources has produced 3055 (33.80 %) of articles. So those 15 journals (sources) were identified this subject research output in core journals (it is showed the core journals at the table 4.34). The second zone group of 101 (8.24 %) journals provides 2968 (32.84 %) articles and the third largest zone of 1109 (90.53 %) of periodicals yield the remaining 3014 (33.35 %) articles. The Bradford multiplier between the number of references in zone 1 and zone 2 is 6.73 while it is 10.98 between zone 2 and zone 3. The average multiplier value is 8.85.

According to *Bradford's* distribution, the relationship between the zone is 1: a: n^2. In contrast is the relationship in each of the present study i.e. 15:101:1109 which fit into *Braflord's* distribution. This shows that core contributions are given by 15 sources, i.e. less than Bradford formulated and the final zones contain a very large number of journals, i.e. much more than the Bradford's formula.

CONCLUSION

The article mainly dealt with the authorship pattern and collaborations in the area of corona virus research output. In the field of science and technology, the studies have conventional the increased growth of multiple authorship and collaborative research. The following facts are derived from this analysis; the most productive year of 2004, 2005, 2006, 2014 and 2013. The publication rate has increased from the value of 0.8 percent to 7.9 percent. It shows that growth rate is increased more than seven times during 1989 to 2014. Collaborative authored publication is highest than the single contribution, only 6.36 percent of articles has single contributions. 93.64 percents of articles were by multi author's productivity. i.e., degree of collaboration value is 0.94.the overall collaborative index value is 5.93. It is essential that in the sphere of Science and Technology large scale studies are carried out to identify the trends in the collaborative research. The three authored production is highest than other authored team. The five authors, ten and above authors contribution is highest than other team. The researcher has identified the author of "Yuen KY**"** is active author.

REFERENCES

1. Amsaveni, N & Vasanthi, R (2012). Mapping of Green Biotechnology Research: A Scientometric Analysis, Journal of Advances in Library and Information Science**,** 1(4) 2012. pp. 181-185.
2. Ajiferuke, I., Q. Burell and J. Tague. 1988. Collaboration coefficient: A single measure of the degree of collaboration in research. *Scientometrics*,.14: 421-433.
3. Amsaveni. N & Vasanthi. R. (2013). Authorship Pattern and Collaborative Research in the Field of Network Security. *Indian Journal of Applied Research*..3 (1): 52-54.
4. Beaver, D. De B. and R. Rosen. 1979. Studies in scientific collaboration 3.Professionalisation and the natural history of co-authorship. *Scientometrics*, 1: 231-245.
5. Clark, B. 1964. Multiple authorship trends in scientific papers. *Science*, Vol.143:820-824.
6. Gupta, B.M., S. Kumar, S. and R. Rousseau. 1998. Applicability of selected probability distributions to the number of authors per article in theoretical population genetics. *Scientometrics*, 42: 325-333.
7. Haitun, S. D. 1982.Stationary scientometric distributions. Part 2. Non-Gaussian nature of scientific activities. *Scientometrics*, .4: 89-104.