

Distributions of Multiple Authors in Two Journals (JASIST and Scientometrics) and in the subject Software Studies: A Case Study

Ravichandra Rao I.K., Professor and Head, Documentation Research and Training Centre, Indian Statistical Institute, 8th Mile Mysore Road, Bangalore 560059, India
E-mail: ikrrao@hotmail.com

and

Sahoo Bibhuti Bhusan, Assistant Librarian, Central Library, INDEST - AICTE Consortium, Indian Institute of Technology- Delhi, Hauz Khas, New Delhi – 110016
E-mail : bibhuti@library.iitd.ac.in

2008, May 29

[Data on distributions of multiple authors in two journals (-- JASIST and Scientometrics) in the subject “software studies” were collected and analyzed. The number of authors per paper in both the journals are increasing; it therefore indicates that the trend is towards collaboration with more than two authors. A similar observation was also made in the area of software studies; i.e., the collaborating nature of research is gradually shifting from 2-authored papers to more than two authors per paper. Further, it has been observed that the distributions of authors confirm to a Poisson distribution.]

1. Introduction

Price¹⁷ in 1963 asserted, “since the turn of century the proportion of the multi authored papers has accelerated steadily and powerfully and it is now so large that if it continues at present rate, by 1980 the single paper will be extinct”. In 1910 the proportion of single authored papers stood 95% and since then the overall decline has been fairly gradual in astronomy as compared with chemistry.

Fields	% of papers having single authors in the year.				
	1910	1934	1940	1960	1963
Astronomy	95	89	86	73	68
Chemistry	82	72	67	40	30

Source: JG O'Connor, 1969, DRTC Seminar 7, p.463

Cherno Gorenko and Muchnik³ studied the subject phosphides and observed the increasing trend of numbers of two and three authored papers, between 1961 and 1973. The number of single author papers has decreased to 21 % in 1973 from 70% in 1961. Two and three authored papers were increased from 15% to 30% and 37% respectively.

Beaver¹ in his article entitled “Does collaborative research have greater epistemic authority?”, pointed out that collaborative research processes have greater epistemic authority than research produced by single researcher. In his study, he argued that single authored papers are slightly more likely never to be cited than collaborative. His data suggests that collaborative research produces significantly more authoritative research.” Kundra¹¹ observed in his thesis an increasing trend towards collaboration and large disparities in the rates of collaboration between ‘pure’ and ‘applied’ areas. He investigated collaboration influences in basic and applied fields of medical Sciences in India. His findings clearly showed a higher degree of collaboration, i.e., an average of 21.95 papers in basic fields of medical science in contrast to an average of 7.63 papers per year in applied field medicine. It is in contrast to a general belief that applied field of medicine is more collaborative in nature than the basic research in medicine. This however was in line with the findings of scholars who also reported a higher degree of international collaboration in basic sciences than in applied field. Lal and Ray¹² have analyzed the data from 511 papers published in four leading journals during 1987-88 in the area of horticulture. They used 5” × 3” slip to record the bibliographic data and then arranged them country-wise. The analysis revealed that multi author papers are more than those of single authors. Single authors’ contribution is 17.42%, whereas double authors contributed 35.02%. It shows that the concept of solo research has been replaced by team research, particularly in the field of applied science like horticulture.

Murphy¹⁶ has analyzed the data from a journal -- “Technology and Culture” -- a quarterly specializing in the history of technology, for a period from 1969-1989. He considered only research articles and excluded reviews, communications, research notes, bibliographies and other occasional features. In his study, he observed that there were 130 papers by single authors, 33 by two authors, 24 by three authors, 8 by four authors and 5 by five authors. The tabulation essentially verifies that Lotka’s law applies to this field within the humanities.

Mendez and others¹⁵ presented an analysis of Spanish bibliography in the field of Physics and Engineering which was retrieved from the database INSPEC and COMPENDEX in the time period 1978-1983. A total of 4,367 documents corresponding to physics and 1,914 corresponding to engineering were analyzed. It has been shown that

literature published by universities follows an exponential curve, while those produced by the research and other organizations, including industries follow a linear growth. The scientific productivity of industry and research organizations is decreasing in the engineering field. They have examined three aspects of co authorship index:

- 1) The habits of co authorship related to the institutions the authors belonged to
- 2) The changes in the co authorship index with time
- 3) The difference in the index when publishing in foreign or Spanish journals.

Research carried out in the more traditional research centers had a bigger co-authorship index. It has also been seen that it is bigger in foreign than in Spanish papers and was growing with time. All these observations seem to point to a trend towards a better quality. Guan and Ma⁶ studied the research performance in the computer science for four major western countries, and India and China on the basis of the INSPEC database during the period of 1993-2002. He observed that scientists in the western countries preferred to work separately or only with one collaborator. Asian scientists preferred “multi-authors” (-- three and four authors) and “mega-authors” (-- five or more authors) collaboration patterns. Suresh Kumar and Garg²⁰ analyzed 2,058 papers published by Chinese authors and 2,678 papers of Indian authors in the field of computer science during 1971-2000; India’s output is significantly higher than the Chinese output. The proportion of a single author papers in China is greater than India’s contribution. The proportion of two author papers, multi-author, and mega-author papers are greater for India as compared to China. The team research in India is greater than in china.

Vijay²¹ Highlights the collaborative research and authorship trends in the area of food science and technology in India he examined and analyzing the authorship patterns in food science and technology periodicals, to study (i) the proportion of single authored articles compared with multi-authored ones, (ii) determine the degree of collaboration in food science and technology, and (iii) to studying the publication pattern of authors from different organizations. The results indicated that that collaborative research was preferred to solo research in the area of food science in India and the degree of collaboration was found to be 0.91. The average number of authors/paper also showed an upward trend from 4.89 in 1994 to 8.2 in 2003.

Hartley⁷ assessed the refereeing times taken from submission to acceptance for psychology journal articles. Based on the data, the author has observed (i) That the papers of single authors were refereed faster than were those of pairs or larger groups, and (ii) That the papers of single authors that contained portions that had previously been delivered at conferences were refereed fastest of all. Bhattacharya and Meyer² argued that co-invention analysis may have the basic concept with co-authorship analysis but Co-invention data exhibits different characteristics. They have observed that

- The level of collaboration in technology has not reached the level of science.
- Co-invention networks do not resemble Co-authorship networks, they are more dyadic in nature and there by characterized by few connections only.

Kretschmer & others¹⁰ examined the similarities between scientists and Web indicators of collaboration, if any for possible use in technology, and policy making. In her study she examined Web hyperlinks and Web visibility indicators, to establish their usefulness, indicators of collaboration and to explore whether similarities exist between Web-based structures and bibliographic structures. Her study indicates that hyperlink structure at the Web don't reflect collaborative structures as appears in bibliographic data. She argued that Web visibility indicators of collaboration are different from hyperlinks and can be successfully used as Web indicators of collaboration. Kim⁹ observed, "The proportion of internationally collaborated papers did not increase substantially during the last two decades, which is against the results of other studies". Based on the analysis of the number of participating, countries, Kim showed that, "multilateral collaboration has increased considerably in the last 20 years, though the proportion of international research collaboration remains stable." Lundberg and others¹⁴ studied the University-Industry collaboration, based on the co-authorship data. Their results show that one-third of the companies (out of 436 companies) that have provided funding to the Universities had not Co-authored any publications with the University. Further the funding indicator identified only 16% of the companies that had co-authored publications. They also observed a case of conflicting trends between funding and co-authorship indicators. Yoshikane and others²² compared the co-authorship networks of the two sub domains in computer science (theoretical and application areas). They argued

1. The degree of importance as the leader, and
2. That as the followers, for each researcher in the co-authorship networks.

They have observed a negative correlation between (1) and (2) is greater in the application area, i.e., the two roles (leader & follower) are separated from each other in the application are more clear than the theoretical area. Ivanisevic and Sapunar⁸ analyzed the number of co-authors in articles published by authors affiliated with domestic (Croatian) and foreign (non- Croatian) institutions in the Croatian Medical Journal. Their study was based on 761 articles published in Croatian Medical Journal between 1992 and 2003. Li-chun and others¹³ argued that co-authorship among scientists represents a prototype of a social network. By mapping the graph containing all relevant publications of members in an international collaboration network: COLLNET. The authors have inferred the structural mechanisms that govern the topology of the social system. Further they argued that the structure of the network affects the

1. Information available to individuals as their opportunities to collaborate.
2. Overall flow of information, and the nature of the scientific community.

They discussed various measures for both the macro-(whole-network) and micro-(actor-centered) structure of collaboration, and then applied to COLLNET. They observed that the COLLNET community displays many aspects of a "small-world," and is somewhat vulnerable to disruption should major figures become inactive.

1.1 Objectives of the Study

Many have worked in the area of “analyses of multiple authorship”. From the above cited studies, one finds a general consensus on the increasing frequency of co-authorship in most disciplines but individual areas of study have been found to vary from one another in their own characteristics rates. The results of studies on “distributions of authors” thus vary considerably from each other. The objectives of this paper are to study the:

- i) Trend in multiple authorship – in two journals (JASIST, Scientometrics) and in the subject “software studies”
- ii) Inter-country and inter-institutions collaboration.
- iii) Trend of citations to the same journals, to other journals and to other types of documents.
- iv) Distribution of multiple authors in the field of software studies.

2. Data Analyses

Data on multiple authors were collected from JASIST and Scientometrics for 12 years, from 1992 to 2004 (Tables 1 and 2). The average number of authors per paper in JASIST varies from 1.45 in 1992 to 2.29 in 2004. The average number of authors per paper in Scientometrics varies from 1.43 in 1992 to 2.35 in 2004. Figure-1 clearly indicates that the average number of authors per paper is increasing, over a period of time - - i.e. collaboration between the authors, may be between the institutions or countries is increasing. For 12 years period, the average of the average number of authors per paper is 1.94 (\bar{x}_1) and 1.88 (\bar{x}_2) for papers in JASIST and Scientometrics respectively. Further t-test (t-value is 0.6426; t_α (for $\alpha = 0.25$) = 2.07) indicate that there is no difference between means of the two populations - - in JASIST and Scientometrics.; Kolmogorov Smirnov ($D_{\max} : 0.0079$ where D_α (for $\alpha = 0.05$) = 0.2776) test further suggests that both the data come from a same “population;” i.e., the distribution of multiple authors in JASIST and Scientometrics are likely to be same!

Table-1. Multiple-authorship data from JASIST

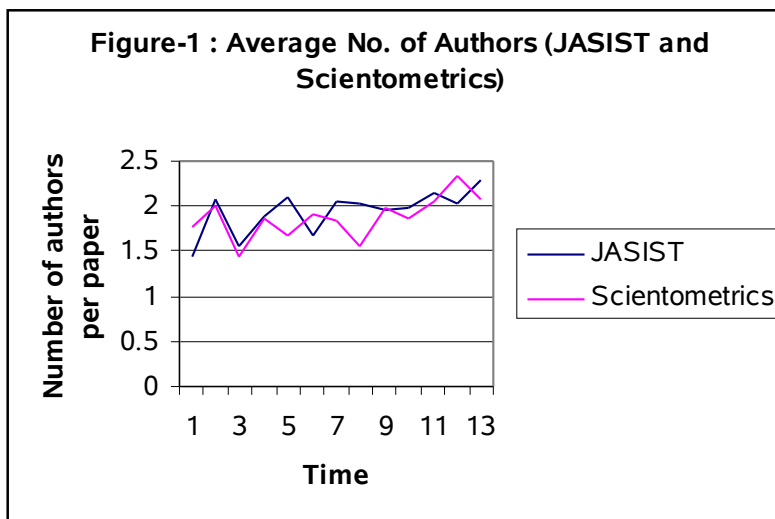
x	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
1	48	18	51	16	39	48	31	34	45	45	52	34	37
2	22	24	20	14	27	38	30	25	36	37	23	46	34
3	2	4	2	3	11	8	18	17	11	14	18	20	13
4	1	2	4	1	4	4	5	7	6	9	6	5	12
5	1	3		1	5		1		3	4	5	2	6
6				1	2				2		3		3
7		1			1		1				2		
8			1					1					
Total	74	52	78	36	89	98	86	84	103	109	109	107	105
mean	1.45	2.08	1.55	1.89	2.09	1.67	2.06	2.04	1.95	1.99	2.14	2.02	2.29
variance	0.54	1.57	1.14	1.32	1.88	0.63	1.19	1.37	1.35	1.2	2.16	0.86	1.82
St. Dev.	0.74	1.25	1.07	1.15	1.37	0.79	1.09	1.17	1.16	1.1	1.47	0.93	1.35

x = is the number of authors in a paper

Table-2: Multiple-authorship data from Scientometrics

x	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
1	39	24	66	39	15	21	40	65	37	43	36	23	33
2	24	19	19	17	10	19	26	18	19	32	28	26	34
3	8	16	4	12	1	10	9	12	23	11	14	26	13
4	5	4	3	4	1	3	3	3	1	5	5	6	10
5	1		1	3	1		4	1	1	3	3	1	
6											2	1	2
9									1			1	
Total	77	63	93	75	28	53	82	99	82	94	88	84	92
mean	1.77	2	1.43	1.87	1.679	1.91	1.84	1.56	1.98	1.86	2.06	2.35	2.09
variance	0.93	0.89	0.65	1.24	0.932	0.8	1.16	0.79	1.49	1.06	1.46	1.63	1.28
St. Dev.	0.97	0.94	0.81	1.11	0.966	0.9	1.08	0.89	1.22	1.03	1.21	1.28	1.13

x = the number of authors in a paper



3. Inter-country and Inter-institutions Analyses

3.1 Inter-country Analyses

The number of papers with authors from different countries (in JASIST) declined from 46.15% in 1992 to 39.71% in 2004 (Table-3) Figure 2 indicates a clear decline in number of papers with authors from different countries. i.e. publications based on collaboration between two countries in these two journals are decreasing considerably. In fact, in scientometrics, it has declined from 39.22% to 21.20 in 2002 (Table-4) In 2004, it was 28.33

Table-3: Multiple-authorship data -- Country-wise (JASIST)

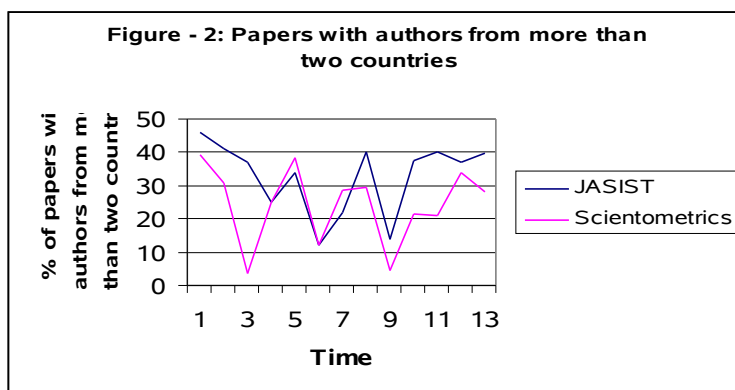
y	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
1	14	20	17	15	33	44	43	30	50	40	34	46	41
2	7	14	8	5	14	6	9	19	6	20	22	22	24
≥2	5		2		3		3	1	2	4	1	5	3
Total	26	34	27	20	50	50	55	50	58	64	57	73	68
% (>= 2)	46.2	41.2	37.04	25	34	12	21.82	40	13.8	37.5	40.4	37	39.71

y : Number of Countries

Table-4: Multiple-authorship data -- Country-wise (Scientometrics)

y	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	
1	31	27	26	27	8	29	30	24	43	40	41	45	43	
2	11	9	1	7	4	4	10	10	2	8	8	13	13	
≥2	9	3		2	1		2			3	3	5	4	
Total	51	39	27	36	13	33	42	34	45	51	52	53	60	
% (>= 2)	39.2	30.8	3.7	25	38.4	6	12.1	28.6	29.4	4.44	21.6	21.2	34	28.3

y : Number of Countries



3.2 Inter-Institution Analyses

We may observe from Tables-5 and 6 that the percentage of authors from different institutions have increased from

- 48.48% in 1992 to 52.94 % in 2004; in fact it was 56.16 % in 2002 and 56.36 in 1996 – the trend is increasing – in JASIST given in Table-5.
- 60.61% in 1992 to 76.47 % in 2004 in Scientometrics given in Table-6.

The Figure-3 indicates that the percentage of authors was centered on 50 to 55% in Scientometrics. However, in JASIST, there is a considerable variation.

Table-5: Multiple-authorship data -- Institution-wise (JASIST)

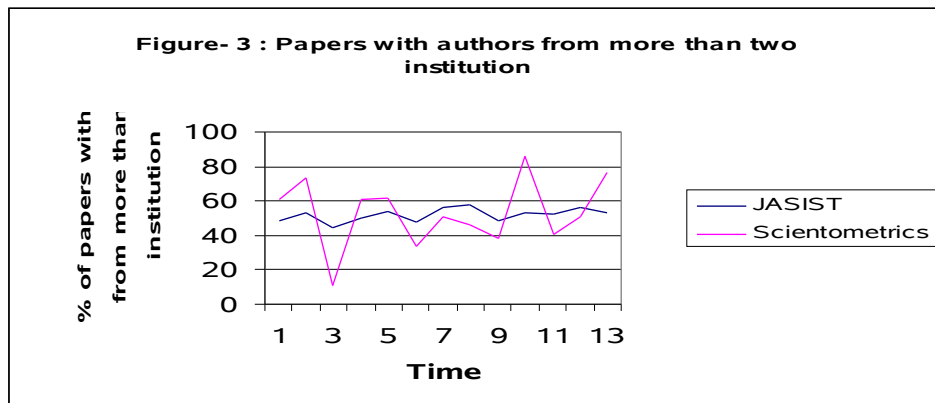
z	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
1	17	16	15	10	23	25	24	21	30	30	27	32	32
2	9	14	10	9	21	20	22	23	24	26	25	33	26
≥ 2	7	4	2	1	6	3	9	6	4	8	5	8	10
Total	33	34	27	20	50	48	55	50	58	64	57	73	68
% (≥ 2)	48.5	52.9	44.44	50	54	47.9	56.36	58	48.3	53.1	52.6	56.2	52.94

z : Number of Institutions

Table-6: Multiple-authorship data -- Institution-wise (Scientometrics)

z	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
1	31	14	24	14	5	17	14	11	23	11	29	26	8
2	11	14	3	17	7	13	22	18	16	23	17	27	30
≥ 2	9	11		5	1	3	6	5	6	21	6	10	22
Total	33	34	27	36	13	48	55	50	58	51	57	73	68
% (≥ 2)	60.6	73.5	11.1	61.1	61.54	33.3	50.9	46	37.9	86.3	40.4	50.7	76.5

z : Number of Institutions



4. Analyses of References

From Tables 7 and 8, we may observe that numbers of

- a) Citations to articles in JASIST by JASIST varies from 5.99% in 1992 to 9.63% in 2004 (Table-7). Where as, in scientometrics, it varies from 2.42% in 1992 to 28.68% in 2004 (Table-8).
- b) Citations to other journals by JASIST vary from 32.58 % in 1992 to 30.32% in 2004; in fact the trend is close to “declining”. A similar trend is also observed even in Scientometrics.
- c) Citations to other documents are centered around (on an average) 45% to 50% in JASIST. In Scientometrics, the figures are 56.53% in 1992 and 37.9% in 2004.
- d) The Figures 4 and 5 indicate the there is hardly any difference in the citation patterns in two journals.

Table-7: Analysis of reference (JASIST)

JASIST	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
to same journal citation to	156	159	143	193	184	288	192	194	321	518	336	358	350
other J citation to	849	449	908	1226	915	1176	1506	824	1409	1378	1320	1641	1102
other D	1601	561	1016	548	1083	1766	1003	1443	1792	1294	1815	1650	2182
Total	2606	1169	2067	1967	2182	3230	2701	2461	3522	3190	3471	3649	3634
% self	5.99	13.6	6.918	9.81	8.433	8.92	7.108	7.88	9.11	16.2	9.68	9.81	9.631
% other J	32.6	38.4	43.93	62.3	41.93	36.4	55.76	33.5	40	43.2	38	45	30.32
% other D	61.4	48	49.15	27.9	49.63	54.7	37.13	58.6	50.9	40.6	52.3	45.2	60.04

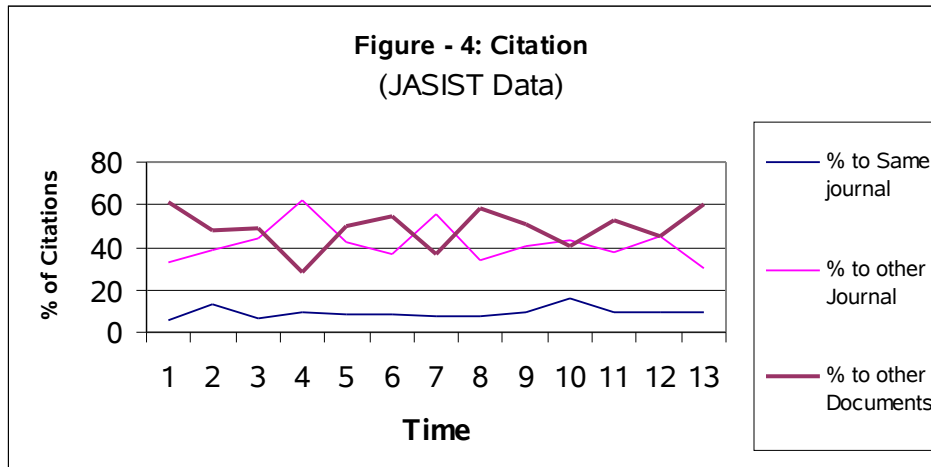
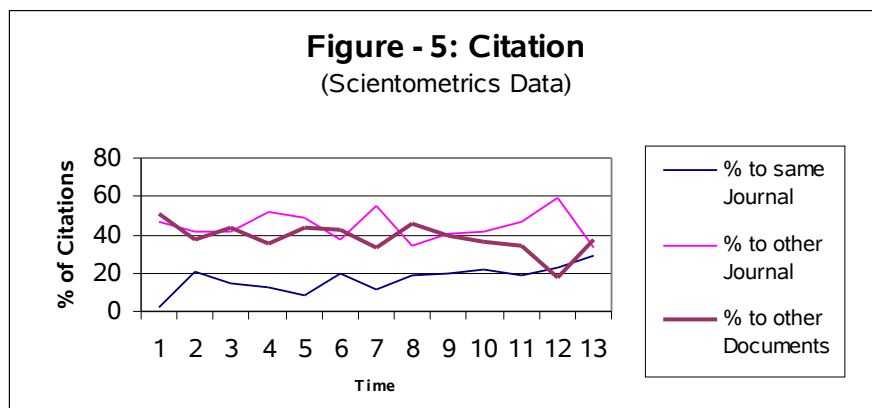


Table-8: Analysis of references (Scientometrics)

Scientometrics	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
self Citation	25	223	171	152	30	180	199	209	324	421	334	504	697
citation to other J	486	451	482	649	178	330	923	378	689	795	854	1309	812
citation to other D	522	406	502	438	160	384	557	503	669	692	629	401	921
Total	1033	1080	1155	1239	368	894	1679	1090	1682	1908	1817	2214	2430
% self	2.42	20.6	14.8	12.3	8.152	20.1	11.9	19.2	19.3	22.1	18.4	22.8	28.7
% other J	47	41.8	41.7	52.4	48.37	36.9	55	34.7	41	41.7	47	59.1	33.4
% other D	50.5	37.6	43.5	35.4	43.48	43	33.2	46.1	39.8	36.3	34.6	18.1	37.9



5. Multi-authorship in software studies

In this section, attempts were made to analyze the percentages of multiple authors (both at India and global level, for a period of 15 years) in software studies. Data were collected from two bibliographic databases – COMPENDEX and INSPEC for a period of 15 years from 1989 to 2003 (Sahoo¹⁹). Tables 9 and 10 give the distributions of multiple

authors for data on India. Tables 11 and 12 give the distributions of multiple authors for the world data. Table 13 gives the average number of authors per paper in software studies for the period 1989 to 2003, both for India and for world. Attempts were also made to identify a suitable probability distribution, based on goodness-of-fit test. For this purpose, a few of the probability distributions, as given below, were fitted to the empirical data, which were given in Tables 9 and 10. The results of the goodness-of-fit test are given in Tables 14 and 15.

Probability Distribution Functions

Sl.	Distribution Functions	Parameters
1.	Poisson Distribution	$P(y) = \frac{e^{-\lambda} \lambda^y}{y!}$ $y = 1, 2, 3, \dots \text{ and } \lambda \geq 1$ $\mu_y = \lambda + 1 \text{ and } \sigma_y^2 = \lambda$
2.	Negative Binomial Distribution	$P(y) = \frac{(k+y-2)!}{(y-1)!(k-1)!} p^k (1-p)^{y-1}$ $y = 1, 2, 3, \dots$ $\mu_y = \frac{kq}{p} + 1 \text{ and } \sigma_y^2 = \frac{kq}{p^2}$
3.	Geometric Distribution	$P(x) = pq^{x-1}$ $x = 1, 2, 3, \dots$ $\text{Mean} = 1/p \text{ and } \text{Variance} = q/p^2$
4.	Lotka's Distribution	$Y_x = \frac{k}{x^\alpha} \quad \alpha > 0 \quad x = 1, 2, 3, \dots$ <p>For $\alpha = 2$ and if p_x refers to relative frequency, then $k = 6/\pi^2$. Thus</p> $Y_x = \frac{6}{\pi^2 x^2}$

Distributions of papers based on the fractional counting are very irregular. Egghe and Rao⁵ developed a methodology (based on an earlier work by Egghe⁴) to compute the theoretical values for distributions of papers. In this methodology, they have assumed that both the distributions of papers and authors follow Lotka's law. However, in this study, it has been observed that the distribution of authors does not follow Lotka's law. Suggested that, in the methodology developed by Egghe and Rao⁵, one may assume that distribution of authors follows a Poisson distribution. In fact in a pilot study (by Rao, Sahoo and Eghge¹⁸) it has been shown that it gives a good result.

Table 9. Percentage of Contributions: Author wise analysis – INDIA From 1989 to 1996								
	1989	1990	1991	1992	1993	1994	1995	1996
x	f(x)	f(x)	f(x)	f(x)	f(x)	f(x)	f(x)	f(x)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	85 (16%)	75 (18%)	61 (16%)	32 (12%)	59 (15%)	74 (15%)	63 (10%)	75 (13%)
2	243 (47%)	199 (47%)	191 (49%)	121 (47%)	183 (47%)	238 (47%)	305 (50%)	276 (47%)
3	122 (23%)	104 (25%)	96 (24%)	77 (30%)	106 (27%)	143 (28%)	158 (26%)	168 (29%)
4	45 (9%)	30 (7%)	34 (9%)	22 (9%)	26 (7%)	33 (6%)	62 (10%)	48 (8%)
5	15 (3%)	7 (2%)	5	4	12 (3%)	10 (2%)	11	10 (2%)
6	4	5 (1%)	4 (2%)	1 (2%)	3 (1%)	6	7(4%)	3 (1%)
7	5				3	1	2	4 (1%)
8	1 (2%)			1	1 (1%)			2
9			1			1 (2%)		
10								
11								
12								
13						2		
14								
15								1 (1%)
16	1							
Total	521	420	392	258	393	508	609	587

5.1 Analysis of Multiple Authors (In software studies with reference to India)

Tables 9 and 10 indicate that percentage of single authorship paper is decreasing – from 16% in 1989 to 12% in 2003, with one or two variations in the year, 1990, 1994 and 1999. Following trends have been observed from the data set for the years 1989 to 2003:

- i. The percentage of single author papers has decreased from 16 % (in 1989) to 12 % (in 2003).
- ii. From 1989 to 1996 the percentage of 2-authored papers has increased; however since 1996, the percentage of 2-authored papers has decreased. This is perhaps due to the increase in collaborative nature of research in the field where more than two authors are contributing to the research work.
- iii. 3-Authored papers have increased from 25% (1998) to 29% (2003) which shows a 4% growth.
- iv. The percentages of 4-authored papers have increased it was 8% in 1989 and finally it is about 10% in 2003.

- v. 5-authored papers have also increased significantly during this period 2.9% (1989) to 5.1% (2003).

Table 10. Percentage of Contributions: Author wise analysis – INDIA (1997 to 2003)

	1997	1998	1999	2000	2001	2002	2003
x	f(x)	f(x)	f(x)	f(x)	f(x)	f(x)	f(x)
(1)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
1	98 (13%)	74 (12%)	120 (15%)	117 (14%)	120 (13%)	122 (13%)	184 (12%)
2	354 (46%)	266 (43%)	345 (42%)	374 (44%)	350 (39%)	368 (38%)	622 (40%)
3	237 (31%)	185 (30%)	260 (31%)	235 (28%)	268 (30%)	291 (30%)	455 (30%)
4	62 (8%)	65 (11%)	73 (9%)	75 (9%)	105 (12%)	106 (11%)	156 (10%)
5	17 (2%)	10 (2%)	20 (2%)	35 (4%)	33 (4%)	40 (4%)	78 (5%)
6	3	10 (2%)	5	10	11	22	22
7		1	1	4	3	2	6
8		4	1	1 (1%)	3 (2%)	1	7
9			1 (1%)			2 (3%)	1
10				1			3 (3%)
11		1 (1%)					
12		1					
13							
14							
15			1		2	1	2
16						1	
Total	771	617	827	852	895	956	1536

5.2 Analysis of multiple authors (in software studies at global level)

Tables 11 and 12 indicates that there is a constant growth of multi authorship papers in software literature, at the global level. The focus has been shifted from the single authorship to the multiple authorship patterns. There is also a transition in the multiple authorship patterns, where we can observe that 2-authored papers are popular in early 1990's. A detailed observation has been given below:

1. The percentage of single author papers have decreased from 38% (1989) to 17% (2003).
2. 2-authored papers do not show a constant growth as we can observe that during 1989 it was 33% and the percentage increases since 1989 to 1994 up to 36% and afterwards it starts decreasing. Finally in the year 2003 it is about 29%.
3. 3-Authored papers have also increased from 17% (1998) to 25% (2003) which shows a 8% growth

4. The percentage of 4-authored papers has increased from 7% in 1989 to 14% in 2003 – thus indicating that there is an increase in the 4-authored paper.
5. 5-authored papers have increased significantly, from 2.4% (1989) to 6.5% (2003).
6. The papers having more than five authors have also increased significantly.

Some important points have been observed from the Tables 9 to 12. Both at national and international level, single authored papers are decreasing and multi authored papers are increasing. It has been observed that the average growth rate of percentage of single author papers is -0.29% during the last 15 years. The negative number may be due to the heterogeneous nature of the population; the single author papers are becoming less and less both in India and at the global level. In case of multi-authored papers we have also observed that there is a trend of 2-authored papers to more than “two author papers;” i.e. the average growth rate of percentage of 2-authored papers is -1.02% during 15 years. The average growth rate in the percentage of three authored papers is 4.5%. So it clearly indicates that multiple authors are preferred in this field during this period where the papers consist of more than two authors. Finally, both in India and at the global level, multi-authored papers are increasing – this is the trend in software studies.

Table 13 and the Figure 6, below show that there is a steady growth in the average number of authors per paper for 15 years from 1989 to 2003. It has however been observed that there is a mixed growth in India. Initially, it was 2.45 in 1989 and it was reduced to 2.31 in 1990. After 1990, it gradually increased and finally it reached 2.67 at 2003 with some exception in the years 1993, 1996, and 1999. Less number of authors per paper indicates a less collaboration in the research activities; the more number of authors per paper indicate that the collaboration among the researcher increases in the field. It can be concluded that the collaboration among the researchers of India is increasing in the field. For the world literature in the software studies it has been observed that the average number of authors per paper varies from 2.131 (1989) to 2.953 (2003). It shows a constant growth in the collaboration of authors.

Table 13 and Figure 6 show that there is a clear increase in the average number of authors per paper over the period of 15 years, in the field of software studies. This means that as the time progresses the more numbers of researchers are involved in the research and development of the field. This shows that there is a significant growth in the field.

**Table 11. Percentage of Contributions: Author wise analysis – World
(1989-1996)**

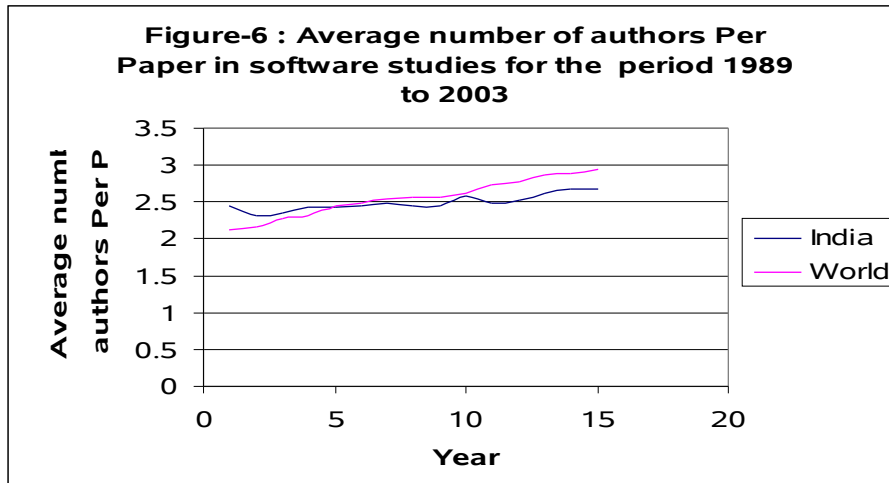
	1989	1990	1991	1992	1993	1994	1995	1996
x	f(x)	f(x)	f(x)	f(x)	f(x)	f(x)	f(x)	f(x)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	17480 (38%)	17952 (36%)	14905 (32%)	9454 (29%)	8348 (24%)	11325 (24%)	12502 (23%)	14104 (23%)
2	15206(33 %)	17200(35 %)	16644(35 %)	11610 (36%)	12707 (37%)	17086 (36%)	19336 (36%)	21878 (35%)
3	7884 (17%)	8769 (18%)	8964 (19%)	6386 (20%)	7641 (22%)	10600 (23%)	12277 (23%)	14599 (23%)
4	3203(7%)	3570(7%)	3747(8%)	2607 (8%)	3354 (10%)	4681 (10%)	5757 (11%)	6754 (11%)
5	1120 (2%)	1293 (3%)	1485 (3%)	1070 (3%)	1273 (4%)	1774 (4%)	2271 (4%)	2613 (4%)
6	526(1%)	576(1%)	637(1%)	488 (2%)	556 (2%)	762 (2%)	1006 (2%)	1160 (2%)
7	233 (1%)	255	297(1%)	198 (1%)	258	340	404	545
8	122	120	132	92	120	191	207	245
9	60	70	71	62	65	100	133	147
10	35	46	52	25	39	58	62	68
11	27	8	27	23	14 (1%)	35 (1%)	48	53
12	16	9 (1%)	14	13 (1%)	12	28	32	36
13	11	14	8	8	5	25	31	24
14	8	10	5	6	6	25	13	21
15	10	2	6	5	33	52	16 (1%)	20
16	15	11	3	20			52	45 (1%)
17	6	9	11 (1%)					
18								1
19								
20								1
21	1 (1%)							
22							1	4
23	1		1					
24								
25			1					
26			1					
27								4
45	1							
Tota l	45965	49821	46985	32049	34431	47082	54148	62322

**Table 12. Percentage of Contributions: Author wise analysis – INDIA
(1997-2003)**

	1997	1998	1999	2000	2001	2002	2003
x	f(x)	f(x)	f(x)	f(x)	f(x)	f(x)	f(x)
(1)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
1	19085 (23%)	14643 (22%)	13684 (19%)	14888 (18%)	15060 (17%)	14419 (17%)	19944 (18%)
2	28137 (34%)	22400 (34%)	23700 (33%)	26737 (33%)	26894 (31%)	25354 (30%)	32450 (29%)
3	19433 (23%)	15991 (24%)	18244 (25%)	21052 (26%)	22264 (26%)	21778 (26%)	28095 (25%)
4	9140 (11%)	7717 (12%)	8956 (12%)	10777 (13%)	11956 (14%)	11841 (14%)	16407 (15%)
5	3639 (4%)	3119 (5%)	3781 (5%)	4471 (5%)	5039 (6%)	5020 (6%)	7268 (6%)
6	1667 (2%)	1386 (2%)	1652 (2%)	1976 (2%)	2333 (3%)	2258 (3%)	3483 (3%)
7	708 (1%)	569	740 (1%)	897 (1%)	1120 (1%)	1073 (1%)	1628 (1%)
8	357	328	416 (1%)	450 (1%)	616	622	943
9	207	199	245	281	318	333	519
10	102	108	133	134	213	185	320
11	71	64	75	86	127	119	199
12	39	36	43	61	91	77	136
13	42	26	37	43	69	61	104
14	25	23	31	25	42	51	78
15	73	20	24	24	33	187	310 (3%)
16	10	63	81	94	218	28	50
17			1	4		4 (3%)	1
18	2 (2%)	4 (2%)	3	4	1	2	
19			1	3 (1%)	1	1	
20		2	2				
21	1	2	2 (2%)		1 (2%)	3	1
22			2			1	
23		1					
24	1						
25		1		1		2	
26		1	1			1	
27	1						
29				3	1		
32		1					
36				2			
37		1					
39	1						
45							
48			1				
66						2 (3%)	
72					1		
112				1			
Total	82741	66705	71855	82077	86398	83422	111936

Table 13: Average number of authors Per Paper in software studies for the period 1989 to 2003

Year	Number of authors per paper		Year	Number of authors per paper		Year	Number of authors per paper	
	India	World		India	World		India	World
1989	2.447	2.131	1994	2.439	2.486	1999	2.48	2.737
1990	2.31	2.152	1995	2.486	2.536	2000	2.52	2.775
1991	2.3597	2.267	1996	2.449	2.557	2001	2.626	2.877
1992	2.43	2.322	1997	2.4428	2.57	2002	2.68	2.892
1993	2.4275	2.454	1998	2.58	2.625	2003	2.6732	2.953



5.3 Distribution of Authors: goodness-of-fit tests

In order to identify a probability distribution function, attempts were made to fit several probability distributions (Poisson, Geometric, Lotka, and Negative Binomial) to the empirical distributions of both India and world data. The results of the goodness-of-fits test are given in Tables 14 and 15. The KS-statistics for the Indian data suggests that Poisson distribution is a close approximation for the Indian data. It also indicates that Poisson distribution is a close approximation for the world data, except for the year 2003. For the year 2003, Negative binomial fits very well.

Thus, in this study we have observed that

- (i). Poisson distribution can be a suitable model to explain the distribution of multiple authorship.
- (ii). Negative binomial is the next best model to explain the distribution of multiple authors.

(iii).Lotka’s model certainly does not fit well – is not a good approximation

Table14: $D\alpha$ and D_{max} – Multiple authorship - India - - KS Test

Year	$D\alpha$	D_{max}			
		Poisson	Geometric	Lotka	Negative binomial
(1)	(2)	(3)	(4)	(5)	(6)
1989	0.0595	0.1260	0.7368	0.4857	0.1356
1990	0.0663	0.1197	-	0.4924	-
1991	0.0686	0.1377	-	0.5036	-
1992	0.0846	0.1279	-	0.5412	-
1993	0.0686	0.1221	-	0.5010	-
1994	0.0603	0.1272	0.8267	0.5048	0.1299
1995	0.0551	0.1636	-	0.6788	-
1996	0.0561	0.1308	-	0.5264	-
1997	0.0489	0.1153	-	0.5433	-
1998	0.0547	0.1064	-	0.5243	-
1999	0.0472	0.0819	-	0.5001	-
2000	0.0465	0.1071	-	0.5126	-
2001	0.0454	0.0715	-	0.5158	-
2002	0.0439	0.0711	0.8368	0.5143	0.0728
2003	0.0347	0.0907	0.8678	0.5218	0.0915

Note: - : not computed, because $\mu > \sigma^2$.

Table-15: $D\alpha$ and D_{max} - World Literature for Multiple authorship distribution - KS Test

Year	$D\alpha$	D_{max}			
		Poisson	Geometric	Lotka	Negative binomial
(1)	(2)	(3)	(4)	(5)	(6)
1989	0.0063	0.0574	0.2226	0.2462	0.0448
1990	0.0060	0.0444	0.3019	0.2694	0.0444
1991	0.0062	0.0354	0.3550	0.3092	0.0507
1992	0.0075	0.0342	0.3770	0.3361	0.0595
1993	0.0073	0.0293	0.4882	0.3902	0.0609
1994	0.0062	0.0266	0.4402	0.3921	0.0637
1995	0.0058	0.0271	0.4462	0.3989	0.0608
1996	0.0054	0.0240	0.4580	0.4011	0.0548
1997	0.0047	0.0237	0.4502	0.3949	0.0452
1998	0.0052	0.0250	0.4355	0.4032	0.0516
1999	0.0050	0.0291	0.4718	0.4329	0.0551
2000	0.0047	0.0265	0.4564	0.4418	0.0636
2001	0.0046	0.0302	0.4274	0.4513	0.0625
2002	0.0047	0.0282	0.4379	0.4499	0.0586
2003	0.0040	0.0363	0.4127	0.4504	0.0485

7. Findings

Following conclusions may be drawn from the current study:

1. The average number of authors per paper in JASIST varies from 1.45 in 1992 to 2.29 in 2004.
2. The average number of authors per paper in Scientometrics varies from 1.43 in 1992 to 2.35 in 2004.
3. For 12 years period, the average of the average number of authors per paper is 1.94 (\bar{x}_1) and 1.88 (\bar{x}_2) for papers in JASIST and Scientometrics respectively.
4. There is no difference between means of the two populations - - in JASIST and Scientometrics.
5. Kolmogorov Smirnov test suggests that both the data come from a same "population"
6. The number of papers with authors from different countries (in JASIST) declined from 46.15% in 1992 to 39.71% in 2004, in fact the figures were 12% and 13.8% in 1997 and 2000 respectively.
7. There is a clear decline in number of papers with authors from different countries. i.e. publications based collaboration between two countries in these two journals is decreasing considerably.
8. The percentage of authors from different institutions has increased from the both cases.
9. Percentage of authors was centered on 50 to 55% in Scientometrics. However, in JASIST, there is a considerable variation.
10. Self-citations to JASIST varies from 5.99% in 1992 to 9.63% in 2004 as given in Table-7 and Figure-7. Where as, in scientometrics, it varies from 2.42% in 1992 to 28.68% in 2004
11. Declined citations to other journals in JASIST from 32.58 % in 1992. 30.32% in 2004; in fact the trend is close to "declining". A similar trend is observed even in Scientometrics.
12. Citations to other documents are centered around (on an average) 45% to 50% in JASIST. In Scientometrics, the figures are 56.53% in 1992 and 37.9% in 2004.
13. There is an increase in the average number of authors per paper over the period of 15 years in the field of software studies for the world literature (2.131 in 1989 and 2.953 in 2003).

14. In the case of Indian literature, there is a steady growth in the average authors per papers for the year 15 years (2.447 in 1989 and 2.6732 in 2003).
15. The collaborating nature of research is gradually shifting from 2-authored papers to more than 2-authored papers to more than two authors per paper, both in India at the global level.
16. The distributions of authors confirm to a Poisson distribution. However, in some cases, particularly for the world data, it also confirms to a negative binomial distribution.

8. References

1. Beaver, Donald Deb. (2004). **“Does collaborative research have greater epistemic authority?”** *Scientometrics*; 60 (3), pp.399-408.
2. Bhattacharya, Sujit; Meyer, Martin. (2004). **“Commonalities and differences between scholarly and technical collaboration: an exploration of co-invention and co-authorship analyses.”** *Scientometrics*; 61 (3), pp.443-456.
3. Chernogorenko, V.B and Muchnik, S.V (1981). **“Scientometric Estimation of a present day study on phosphides.”** *Scientometrics*; 3(2), pp. 127-134.
4. Egghe, L (1993). **“Consequences of Lotka’s law in the case of fractional counting of authorship and of first author count.”** *Mathematical and computer modeling*, 18(9), 63-77.
5. Egghe, L; Rao, I K R. (2002). **“Duality revisited: construction of frequency distributions based on two dual Lotka laws.”** *Journal of the American Society for Information Science and Technology*; 53 (10), p.789-801.
6. Guan, J.C and Ma, N. (2004). **“Research performance in Computer Science.”** *Scientometrics*, 61(3), pp. 339-359.
7. Hartley, James. (2005). **“Refereeing and the single author.”** *Journal of Information Science*; 31 (3), pp.251-256.

8. Ivanisevic, Ranka; Sapunar, Damir. (2006). “**Multiple *authorship* in a small medical journal: a case study of the Croatian Medical Journal.**” *Journal of the American Society for Information Science and Technology*, 57(8), pp. 1073-1078.
9. Kim, Ki-Wan. (2006). “**Measuring international research collaboration of peripheral countries: taking the context into consideration.**” *Scientometrics*, 66(2), pp. 231-240.
10. Kretschmer, Hildrun; Kretschmer, Ute; Kretschmer, Theo. (2007). “**Reflection of *co-authorship* networks in the Web: Web hyperlinks versus Web visibility rates.**” *Scientometrics*, 70(2), pp. 519-540.
11. Kundra, Ramesh. (2002). “**Quantitative Analysis of Medical Literature in India: A Bibliometric Study.**” *Thesis submitted to the Dept. of Library and Information Science of the Pune University for the Award of Degree of Doctor of Philosophy in Library and Information.* Guide: I.K.Ravichandra Rao, DRTC, Indian Statistical Institute, Bangalore.
12. Lal, Arjunand Ray, P.K. (1993). “**Pattern of Research Contribution in leading Horticulture Journal of the World: a Comparative study.**” *IASLIC Bulletin*, 36(3), pp. 95-102.
13. Li-chun, Yin and others (2006). “**Connection and stratification in research collaboration: an *analysis* of the COLLNET network.**” *Information Processing & Management*, 42(6), pp. 1599-1613.
14. Lundberg, Jonas and others. ((2006). “**Collaboration uncovered: exploring the adequacy of measuring university-industry collaboration through *co-authorship* and funding.**” *Scientometrics*, 69(3), pp. 575-589.
15. Mendez, A., Fernandez, M.F and Lopez, Aguada G. (1987). “**Six Year Spanish Scientific Activity in Physics and Engineering through INSPEC and COMPENDEX.**” *Scientometrics*, 12(1/2), pp. 81-100.
16. Murphy, L.J. (1973). “**Lotka’s Law in Humanities?**” *Journal of the American Society for Information Science*, 24 (6), pp (161-162).
17. Price, Derek de Solla (1963).”**Little Science, Big Science.**” *New York.1963. p.87-89.*
18. Rao, I. K. Ravichandra; Sahoo, Bibhuti Bhusan and Egghe, L. (2003). “**A Distribution of papers based on Fractional Counting: An Empirical Study.**” *Proceedings of the 9th International Conference on Scientometrics and Informetrics, Dalian University of Technology Press, Dalian, China. p. 241-250.*

19. Sahoo, Bibhuti Bhusan.(2006). “**Scientometric Study of Literature in Software Studies in India with a comparison to the World Literature.**” *Thesis Submitted to the Dept. of Library and Information Science of the Pune University for the Degree of Doctor of Philosophy in Library and Information Science.* Guide: I. K. Ravichandra Rao, Documentation Research and Training Center, Indian Statistical Institute, Bangalore.
20. Suresh Kumar and Garg, K.C. (2005). “**Scientometrics of Computer Science Research in India and China.**” *Scientometrics*,64 (2), pp. 121-132.
21. Vijay, K.R. (2005). “**Bibliometric Study of Research Publication Trends among Indian Food Scientists and Technologists.**” *Annals of Library and Information Studies.* 52(3). 77-81.
22. Yoshikane, F., and others. (2006). “**Comparative *analysis* of *co-authorship* networks considering authors' roles in collaboration: differences between the theoretical and application areas.**” *Scientometrics*, 68(3), pp. 643-655.