International Journal of Library Science;
Volume 09; Issue No. 3; Year 2013;
ISSN 0975-7546; Int. j. lib. sci.

An Evaluative Study on Citation Patterns of Sankhyá

Prabir Kumar Das and **Jiban K Pal**

Library, Documentation and Information Science Division,
Indian Statistical Institute, 203 B. T. Road, Kolkata - 700108, W.B, (India);
E-mail : prabirdas2003@yahoo.com; jiban@isical.ac.in

ABSTRACT

This study analyses 3750 citations appended to 199 peer-reviewed articles published in Sankhya during 2003 to 2007. Critical examinations have been made on average citations occurred in each publication, various source materials cited, highly cited keywords, frequently cited journals, and identifying the core journals in statistical research. Bradfor's scattering of cited journals (zonal distribution) has also been carried out. Findings reveal that an average of 18.84 cited references were appended in each publication, and thereby reinforces the proposition of discipline oriented citation behavior of scholarly literatures. The paper illustrates wide variety of source materials, where journal-articles get cited predominantly and the citation of web-resources is very poor; thus citation behavior of Sankhya exhibits a close resemblance with the usual practice of S&T journals. It suggests an average of five keywords to be transcribed by the authors in each article. The study identifies a total of 2732 journal-citations in 372 unique titles, thus journal citation density is derived as 7.35. In fact top 12 journals have contributed more than 50% citations, subsequently top five (core journals) received more than one-third (34%) of the total journal-citations. Annals of Statistics is the most highly cited journal; followed by JASA, Biometrika, JRSS-B, and Annals of Mathematical Statistics. Bradford's plot (cumulative citations vs. logarithm of journal ranks) presents a deviation of classical S curve. Above all Sankhya could stake claim as one of the most authoritative source of scholarly literatures on statistics and allied areas of research.

Keywords: Citation analysis, collection management, statistics journal, Sankhya, India.

Library of Congress Classification: Z669.8

1. INTRODUCTION

Citation analysis has been widely recognised as a convenient method to evaluate the scientific activities employing various bibliometric indicators on the literatures available in premier communication channels of a particular field of study. There has been a broad array of empirical studies in different domains of natural, social, and behavioural sciences; since Gross brothers had conducted a statistical method of grading the scientific periodicals based on citation counts (*Gross & Gross, 1927*). Further advocacy of citation analysis has been made in 1950's, when Eugene Garfield had mooted the significance of citation indexes in scientific evaluations (*Garfield, 1972*). However, the new understanding of de Solla Price through his landmark contribution stimulated to foster an effort towards mapping the intellectual structure of scientific disciplines (*de Solla Price, 1963*). He pointed out - science is not a unified whole, but a mosaic of specialty areas… Consequently Ziman viewed that a scientific paper does not stand-alone, it is embedded in the 'literature' of a subject (*Ziman, 1968*). In fact, contributions to scientific knowledge are predominantly crystallized in the form of

scientific articles conceiving nascent micro-thoughts, theories, innovations, or new synthesis of existing facts. Thus intellectual links between an individual paper and subsequent literatures are formally made in the form of citations; thereby providing the basis of measuring contributions pertaining to the subject concerned. Therefore, citations analysis identifies the relationship between cited and citing documents, as well as deals with the study of these relationships in multiple dimensions (*Egghe & Rousseau, 1990*).

It is evident that citation studies are based on the semantic relations between two literatures; one that cites, another which is cited. Invariably the citations are appended in scientific articles (often with a list of references), primarily to support the insights drawn in previous works and becomes useful to establish a new point-of-view, so as to make a coalition of intellectuality among the cited and citing works. Therefore the analysis of citation patterns and behaviors of scientific articles appeared in a scholarly journal would enable us to map the cognitive structure of the scientific specialty. Here an attempt has been made to investigate the citation patterns of *Sankhya* based on the literatures appeared in this scholarly journal during 2003 to 2007.

The motivation of the present study lies in the scarcity of citation literature on this particular specialty, while single-journal citation studies are more frequent in other major disciplines of natural, social and even behavioral sciences. Reportedly such studies are rarely made in the field of statistics. In fact the authors have identified a single occurrence of citation-study on statistics and probability (*Stigler, 1994*); but no such intrinsic effort has been found based on single-journal of this domain. However, no Indian journal of statistics ever been studied from India and abroad; even though, India has a rich tradition of statistical research (*Editorial, 1933*) and the Indian contribution to the field has grown significantly during last few decades. Nonetheless, the citation mapping in statistical research becomes very crucial over other disciplines, since statistics is a universal tool of inductive inference and technological applications.

Sankhya being the first Indian journal of statistics (scholarly and peer-reviewed) with international reputation (by circulation and author's participation) justifies the special significance to conduct a study for representing citation-pattern of this scientific specialty.

2. THE CONCEPT WITH MANY FLABOURS

The phenomenon of 'citation' is very difficult to characterize. Perhaps the greatest potential of it lies in the new insights that correlates previous references as authentic source of information having sort of research value, otherwise to be considered a vague. Reportedly the concept has emerged from the identification of source of an idea, developed after the invention of printing during renaissance; but perhaps the formal use of citation was traced back to 16th century (*White, 1985*). Thereafter many researchers have perceived the citation analysis from numerous angels; may be conflicting, seldom ambiguous. Here we are intended to sketch the landmark views rendered by eminent scholars, rather to attempt an extensive review of citation literatures.

Garfield et al. (1978) viewed the citation analysis as an analytical tool that becomes useful as science indicator. Price (*de Solla Price, 1986*) distinguished referencing with citation and explained the referencing is more backward looking concept while citation is forward looking process. *Sandison* (1989) emphasized that citation is a representation made by the author who likes to show the relation between the documents. *Johnson* (2000) opined that citation studies can be an effective tool to guide collection development in academic libraries. *Small* (2004) noted the citations are symbolic payments of intellectual debt. *Williams & Fletcher* (2006) learned citation analysis as a non-intrusive method of identifying patterns in using research materials. Again, *Ronald Rousseau* (2008) somewhat differently viewed the citation analysis as an effective tool for information retrieval. Practically he described the concept as "a collaborative peer effort to analyze and promote the quality of scholarly publication and research" (*Rousseau & Hu, 2010*). In no doubt, citation has become very useful tool for analyzing the research activity in any scientific specialty, and also useful to library managers for collection development, budget allocation, and many other purposes. Thus researchers as well as research institutions are increasingly utilizing the citation-analysis as reliable and convenient tool for evaluating their research outcomes.

3. SCOPE AND OBJECTIVES

Primary objective of this study is to analyze the citation pattern adopted by *Sankhya* during the period from 2003 to 2007. However, the study is conducted to assess the research impact of the journal in statistical science research. It has been worked out based on the publications (articles) appeared over the issues of this journal; thus, so called less scholarly communications like book-reviews, preface, editorial-notes, letters-to-editor, corrigendum, obituary, etc. are excluded from the purview of this study. The paper analyses 3750 citations appended to 199 peer-reviewed articles covered in the source journal during the study period.

Indeed, specific objectives of the study are;

- to find out the chronological distribution of citations.
- to map the frequency distribution of cited references.
- to enumerate the types of source documents cited.
- to determine the most frequently cited keywords in this channel.
- to ascertain the applicability of Bradford's law of scattering.
- to identify the core journals in this scientific specialty.
- to examine the availability of the core-journals at the Institute library.
- to justify the relevance of the source journal in statistical science research.

4. SOURCE JOURNAL

Sankhya – The Indian Journal of Statistics, an international peer-reviewed scholarly journal published by the Indian Statistical Institute since June 1933. It seems to be the first Indian journal on Statistics

founded with the editorship of Professor Prasanta Chandra Mahalanobis, the father of statistics in India. Actually this journal emerged to extend the unique perception of Professor Mahalanobis toward consideration of statistics as key-technology and to unfold the twin aspects of statistics, both theoretical and applied (*Rudra, 1996*). However it carried much of the path-breaking research works of P. C. Mahalanobis and his close associates like R. C. Bose, S. N. Roy, S. N. Bose, C. R. Rao and J. K. Ghosh.

In pursuance of above philosophy, the journal provides an excellent communication channel for exchanging innovative ideas and developments in different dimensions of statistics, which make Sankhya an effective and reliable representation of current statistical research. It is therefore publishes peer-reviewed articles representing original research in the broad areas of theoretical statistics, probability and applied statistics to pursue vigorous research activities. Thus the journal has played a decisive role to the advancement and dissemination of scholarly information in the statistics and allied areas. In terms of visibility, the articles of this journal are covered in premier indexing and abstracting services viz. Mathematical Reviews (MR), Statistical Theory and Methods Abstract (STMA), Indian Science Abstracts (ISA), Zentralblatt fur Mathematik (ZMATH), Current Index to Statistics (CIS), Google-scholar, Scopus, etc. In fact, Science Citation Index (SCI) used to cover the Journal for a period from 1966 to 1992. Now the publishing giant Springer has signed an agreement (in 2009) to co-publish the *Sankhya* allowing researchers to access through a global platform Springer-Link. Springer enables online submission system along with Editorial Manager; thus allows faster peer-review process for more speedy publication of the articles, and facilitates alerting services through Online-First platform (*Sankhya website*).

The journal publishes quarterly issues usually come out in February, May, August and November. It is indeed essential to mention that this house journal in its' seventy-five years of journey has undergone several changes; by means of splitting into a number of series (A, B, C, even D), and gradually it squeezed into a single-title as appeared in 2003. Further splitting was made in 2008 into two different series with varying scope and ISSN. Naturally in its' metamorphic phases, the journal has created significant queries among the statisticians as well as to bibliometricians. So, an effort has been pursued to catch up the glimpses of *Sankhya* with 'unified title' appeared during 2003 to 2007. Therefore, the citation analysis of peer-reviewed research articles appeared in five volumes (65 to 69, covers twenty issues) of *Sankhya* would certainly be an indicative of contemporary trends of citation in statistical research.

5. METHODOLOGY

In order to address the aforesaid objectives, primary data of the study has been collected from *MathSciNet* (AMS) combining diverse searchable fields. So, bibliographic data of articles having source-journal as Sankhya in the byline and published during 2003 to 2007 were retrieved from the source database. Complete searching followed by necessary data cleaning yielded 199 hit records, which were useful for counting the citations appended to each article consulting physical volumes of the source-journal. A total of 3750 citations appended to 199 articles covered during the study period

were collected and tabulated according to bibliographic-form, citation-frequency, cited-journal, cited-keywords, etc.

Therefore, necessary bibliographic elements of individual citations were compiled manually employing systematic methods and recorded using MS-Excel spread sheets. Subsequently recorded data have been analyzed for making observations and interpretations. Various bibliometric methods and techniques are applied to determine the mean-value of citations per article, scattering of cited-journals, ranking of core journals, etc. Bradford's law of scattering has been used to ascertain the applicability of the Law on journal citation of *Sankhya.*

Thus a thorough analysis of collected data has been worked out and the descriptive research design was adopted.

6. DATA ANALYSIS AND FINDINGS

6.1 Year-wise Distribution of Citations

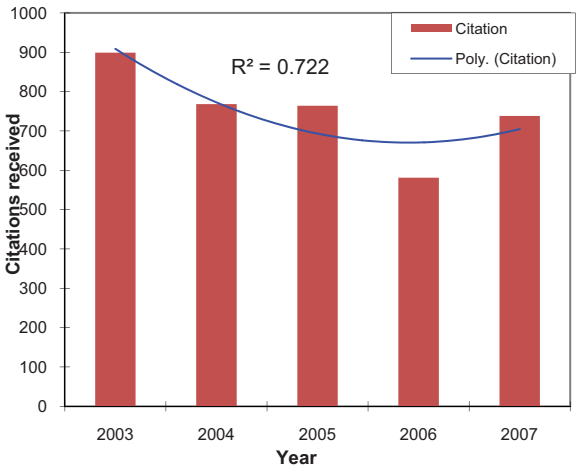
Table 1 shows year-wise or volume-wise distribution of citations (cited references), indicates that 5 volumes (20 issues) appended 3750 citations in 199 citing articles published during 2003 to 2007. It appears that every issue published approximately 10 articles and each article has an average of 18.84 cited references. In fact the numbers of articles and corresponding citations have decreased gradually over the years, except in 2007. Analysis reveals that highest number of citations (899) received in 2003 (23.97%) and least citations appeared in 2006 (15.49%); but the average citation per article (20.75) was maximum in 2006 and minimum in 2004. However a steady trend of citation following second-degree polynomial type (R^2=0.7227) is observed, as presented in Fig.1.

Further analysis depicts that citation pattern of Sankhya conforms to many other scientific disciplines, thus complies the Price's 'norm-of-scholarships' for scientific literatures (*de Solla Price, 1970*). Several literatures of citation studies imply that science and technology related fields (including specialized research areas) usually appended lower number of citations, whereas in the softer fields (like social and behavioral sciences) the number of citations seems to be higher. Therefore this study reinforces the proposition of discipline oriented citation behavior of scholarly literatures (*Zainab & Anyi & Anuar, 2009*).

Table 1. Year-wise distribution of citations

Year	Volume (issue)	No of articles	Ref. cited	Percentage	Avg. citations per article
2003	65 (1-4)	49	899	23.97	18.35
2004	66 (1-4)	45	768	20.48	17.07
2005	67 (1-4)	38	764	20.37	20.11
2006	68 (1-4)	28	581	15.49	20.75
2007	69 (1-4)	39	738	19.68	18.92
Total	Five (twenty)	199	3750	100	18.84

Figure 1. Trend line of citations (Year-wise)



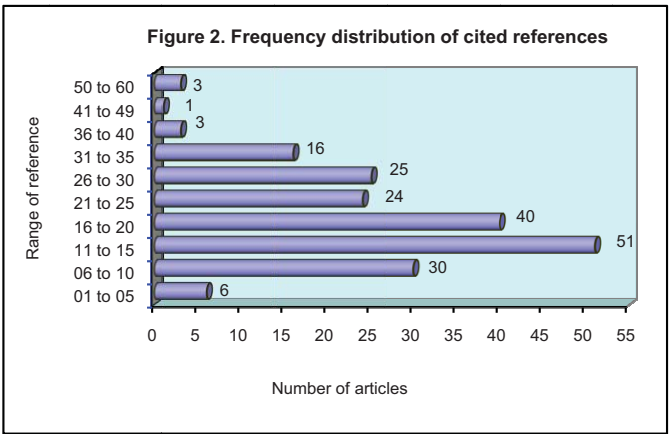
6.2 Frequency Distribution of Citations

Table 2 and adjacent figure depicts the frequency distribution of citations that are appended to each publication of *Sankhya* during the study period. Subsequently, the number of articles (frequency) belonged to a group (range) is tabulated and the statistical mean value of citations has been calculated. Distribution shows that a majority of articles (140) were appended cited-references within the range of 11 to 30; whereas, only 3 articles have appended 50 or more citations. About 19 citations per article (Mean = 18.81) is found during the period of study.

Worthy to note that, highest number of citations (57) received by an article published in 2007 (in vol. 69, issue 3, p.514-547), and least number of citations (3) received by another article published in the same volume (issue 2, p.137-161). However, only 4 out of 199 articles have appended more than 40 cited references. Figure 2 clearly shows the frequency distribution of cited references is asymmetrical or skewed.

Table 2. Mean citations per article in Sankhya

Range of references	Frequency (f)	Mid value (m)	f x m	Mean value
01 to 05	6	3	18	$\overline{X} = \frac{\sum fm}{N}$
06 to 10	30	8	240	
11 to 15	51	13	663	
16 to 20	40	18	720	
21 to 25	24	23	552	
26 to 30	25	28	700	$= \frac{3745}{199}$
31 to 35	16	33	528	
36 to 40	3	38	114	
41 to 49	1	45	45	**=18.81**
50 to 60	3	55	165	
Total	199 (N)		3745	



6.3 Types of Source Materials Cited

Researchers cite various source materials to establish their research findings. Usually choice of cited-references varies with numerous factors – viz. nature of research, discipline of study, type of publication, method of research, availability of the information resources and many others (*Bornmann & Daniel, 2008*). Here the study of citation reveals that contributors have been found to use a wide variety of source materials for reporting their research, as shown in Table 3.

Table 3. Distribution of source materials cited

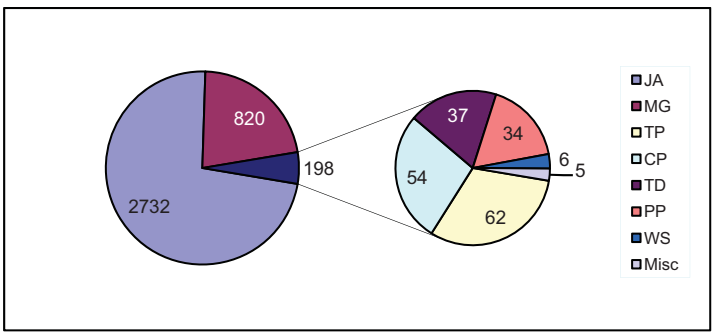
Vol. Part, Year	No. of articles	Source materials used								Ref. cited
		JA	MG	TP	CP	TD	PP	WS	Misc	
65 Pt-1, 2003	12	178	59	5	4	3	2	-	-	251
65 Pt-2, 2003	15	179	73	1	2	3	1	-	-	259
65 Pt-3, 2003	13	189	42	5	3	4	4	-	1	248
65 Pt-4, 2003	9	87	47	1	2	1	1	2	-	141
66 Pt-1, 2004	10	159	41	8	9	1	1	-	-	219
66 Pt-2, 2004	12	109	32	1	2	1	2	-	-	147
66 Pt-3, 2004	12	148	45	3	4	1	1	-	-	202
66 Pt-4, 2004	11	150	40	5	2	1	-	-	2	200
67 Pt-1, 2005	6	66	23	1	1	1	-	-	-	92
67 Pt-2, 2005	13	225	62	6	2	1	3	1	-	300
67 Pt-3, 2005	8	125	34	1	5	3	5	-	-	173
67 Pt-4, 2005	11	145	48	3	2	1	-	-	-	199
68 Pt-1, 2006	7	104	33	3	2	2	2	-	-	146
68 Pt-2, 2006	8	106	55	4	1	1	2	1	-	170
68 Pt-3, 2006	6	105	25	7	2	4	2	-	1	146
68 Pt-4, 2006	7	91	20	-	1	2	4	1	-	119
69 Pt-1, 2007	6	99	19	4	2	1	1	-	-	126
69 Pt-2, 2007	12	106	36	1	4	2	1	-	1	151
69 Pt-3, 2007	9	209	40	-	3	-	2	1	-	255
69 Pt-4, 2007	12	152	46	3	1	4	-	-	-	206
TOTAL	**199**	**2732**	**820**	**62**	**54**	**37**	**34**	**6**	**5**	**3750**

[JA = journal articles. MG = books, chapters of composite books, entries of multi-volume series, etc. TP = technical papers includes working paper, discussion paper, project report, laboratory notebook, etc. CP = conference proceedings. TD = thesis and dissertations. PP = preprints and unpublished manuscripts or mimeo. WS = website links. Misc. = miscellaneous items, includes user manual/ guide, audit reports and annual report].

The citation pertain to various forms of materials like journal articles, monographs, conference papers, preprints, technical papers, thesis and dissertations, etc. The analysis identifies an overwhelming emphasis on citations from the journal literatures account for 72.85% (2732) of total citations; followed by monographs (21.87%), and rest of the citations were covered by technical papers (1.65%), conference proceedings (1.44%), thesis and dissertations (0.99%), and unpublished manuscripts or mimeo (0.91%), etc. Notably, the use of web-resources is likely poor cited (0.16%). However, miscellaneous items comprise of user-manuals or user-guides, audit reports, annual reports became into existent (0.13% only).

It is evident that (Fig.3) alike many other scientific disciplines, journal articles were predominating in citations of Sankhya; thus citation pattern seems to conform usual practices in science and technology journals.

Figure 3. Various source materials cited



6.4 Highly Cited Keywords in *Sankhya*

Table 4. Rank list of author-assigned keywords

Keyword	Frequency	Cu. frequency	Percentage	Cu. %
Asymptotics	12	12	1.238	1.238
Efficiency	9	21	0.929	2.167
Maximum likelihood estimation	9	30	0.929	3.096
Bootstrap	8	38	0.826	3.922
Bayesian statistics	7	45	0.722	4.644
Order statistics	7	52	0.722	5.366
Regression	7	59	0.722	6.089
Consistency	6	65	0.619	6.708
Characterization	5	70	0.516	7.224
Nonparametric regression	5	75	0.516	7.740
Quantile regression	5	80	0.516	8.256
Bayes estimation	4	84	0.413	8.669
Coherence	4	88	0.413	9.082
Exchangeability	4	92	0.413	9.494
Goodness of fit	4	96	0.413	9.907
Model selection	4	100	0.413	10.320
Weak convergence	4	104	0.413	10.733
10 keywords used	Thrice each	134	3.096	13.829
63 keywords used	Twice each	260	13.003	26.832
709 keywords used	Once each	969	73.168	100.000
Total	969		100	100

Keyword is supposed to be one of the best indicators to understand and grasp the thought content of scholarly communications (articles) in any area of research. A total of 969 unique keywords (author-assigned) were identified from 199 articles studied here, thus an average of about 5 keywords were transcribed in each article. Keywords were collected directly from the articles itself, as assigned by the contributing authors to represent the thought content. Thus a large number of keywords were found to cover a wide range of areas of research on statistics and allied sub-domains. A truncated list of keywords with corresponding frequencies is presented in the Table 4.

Analysis shows that 'asymptotics' is the most frequent keyword among the dataset thereby topped in the list, followed by followed by maximum likelihood estimation, bootstrap, bayesian statistics, order statistics, and regression. 63 keywords appeared twice each and 709 keywords appeared only once each. The analysis of key words certainly would help the aspiring authors to ascertain the deeper scope and coverage of Sankhya with greater precision, and to make further strategies in identifying the gaps of research in statistical science. It is clearly observed that the Journal as such no practice of controlling standard vocabulary for assigning keywords.

6.5 Ranking of Cited Journals

Table 5 depicts the rank list of cited-journals as appended in the articles of *Sankhya* during the study period. It has been prepared exclusively on the basis of total citations received by them. A total of 372 unique journal-titles received a total of 2732 citations. The journal citation density has been derived as 7.35, is quite significant. The citation density is an average between the number of references cited and number of journals cited (*Garfield, 2007*).

The study reveals that *Annals of Statistics* is highly cited journal in Sankhya having 11.68% (319) of total journal citations; followed by *Journal of the American Statistical Association* (240), *Biometrika* (154), *Journal of the Royal Statistical Society – B* (118) and *Annals of Mathematical Statistics* (101). However, least citations received by 196 journals having only 1 citation each; thus constituted 7.17% of total cited journals.

Table 5. Rank list of journals based on citations.

Rank	Journal title	Citations	%	Cu. Citation	Cu. %
1	Annals of Statistics	319	11.676	319	11.676
2	Journal of the American Statistical Association	240	8.785	559	20.461
3	Biometrika	154	5.637	713	26.098
4	Journal of the Royal Statistical Society – B	118	4.319	831	30.417
5	Annals of Mathematical Statistics	101	3.697	932	34.114
6	Statistics and Probability Letters	79	2.892	1011	37.006
7	Journal of Statistical Planning and Inference	76	2.782	1087	39.788
8	Journal of Multivariate Analysis	71	2.599	1158	42.387
9	Annals of Probability	63	2.306	1221	44.693
10	Technometrics	59	2.160	1280	46.852
11	Biometrics	58	2.123	1338	48.975
12	Sankhya: The Indian Journal of Statistics - A	52	1.903	1390	50.878

13	Communications in Statistics: Theory & Methods	49	1.794	1439	52.672
14	Journal of Applied Probability	37	1.354	1476	54.026
15	Annals of the Institute of Statistical Mathematics	34	1.245	1510	55.271
16	Canadian Journal of Statistics	33	1.208	1543	56.479
17	Econometrica	32	1.171	1575	57.650
18	Statistica Sinica	31	1.135	1606	58.785
19	Scandinavian Journal of Statistics	29	1.061	1635	59.846
20	Statistical Science	26	0.952	1661	60.798
21	Journal of Nonparametric Statistics	25	0.915	1686	61.713
22	Journal of Econometrics	24	0.878	1710	62.592
23	Probability Theory and Related Fields	23	0.842	1733	63.433
24	Sankhya: The Indian Journal of Statistics	21	0.769	1754	64.202
25	Advances in Applied Probability	20	0.732	1774	64.934
26	Stochastic Process Applications	19	0.695	1793	65.630
27	Computational Statistics and Data Analysis	18	0.659	1811	66.288
28	Statistics in Medicine	17	0.622	1828	66.911
29	Sankhya: The Indian Journal of Statistics - B	16	0.586	1844	67.496
30	International Statistical Review	15	0.549	1859	68.045
31	4 Journals having 14 citations each	56	2.050	1915	70.095
32	3 Journals having 13 citations each	39	1.428	1954	71.523
33	4 Journals having 12 citations each	48	1.757	2002	73.280
34	5 Journals. having 11 citations each	55	2.013	2057	75.293
35	4 Journals having 10 citations each	40	1.464	2097	76.757
36	4 Journals having 9 citations each	36	1.318	2133	78.075
37	3 Journals having 8 citations each	24	0.878	2157	78.953
38	5 Journals having 7 citations each	35	1.281	2192	80.234
39	6 Journals having 6 citations each	36	1.318	2228	81.552
40	8 Journals having 5 citations each	40	1.464	2268	83.016
41	20 Journals having 4 citations each	80	2.928	2348	85.944
42	28 Journals having 3 citations each	84	3.075	2432	89.019
43	52 Journals having 2 citations each	104	3.807	2536	92.826
44	196 Journals having 1 citation each	196	7.174	2732	100
	Total 372 unique titles	2732	100		

6.6 Bradford Scattering of Cited Journals

Table 6 presents a distribution of journals and corresponding number of citations in different three zones, known as Bradford Zones. Bradford law of scattering describes a quantitative relation between the cited journals (references) into three approximately equal zones in order of descending productivity; and successive three zones follow a common pattern, expressed as 1: n: n^2, where *n* is a multiplier. Thus it explains distribution of citations over the journal titles relevant to a given field. The first zone (read as nuclear zone) is highly productive (occurs maximum density) and can be treated as core group of journals that exerts adequate influence in the field. Second zone is moderately productive zone, whereas third zone (as peripheral zone) is low productive, in compare to other two zones.

Here a test was carried out to ascertain whether *Bradford Law of Scattering* is applicable to citation dataset of *Sankhya*. Table shows 5 journals (belonging in nuclear zone) produced 932 citations

(34%), next 25 journals covered 927 citations (33.9%), and rest 342 journals covered 873 citations (32%). According to Bradford, these zones, thus identified will approximately form a geometric series of 1 : n : n^2. In the present study, the ratio of the number of journals in three successive zones are 5 : 25 : 342 (i.e. 1 : 5 : 68). So, the number of journals in the third zone far outnumbers the expected value (125), thus does not fit into Bradford's distribution. However, the value of multiplier (n) resembles with the Bradford's empirical study (*Bradford, 1985*) on the data for geophysics and lubrication, where he suggested the value of 'n' to be 5 as a representative number... So the scattering of journals pursued in this study makes sense of Bradford distribution.

Table 6. Distribution of citations to journal titles

Bradford Zones	No. of journals	Percentage	No of citations	% of Citations
I : Most productive	5	1.34 %	932	34.11 %
II : Moderately productive	25	6.72 %	927	33.93%
III : Least productive	342	91.94 %	873	31.96%
Total	372	100 %	2732	100 %

The classical Bradford's plot usually assumes the shape of an elongated S, where the core journals are those whose points lay on the initial curved part of the graph until it tangentially becomes a straight line, and the sloping part at the top of the Bradford curve is called the groos droop. If the distribution conforms to Bradford's law, then the graph is known as *Bradford bibliography* that should display three distinct characteristics of the regions - i) a rapid rise for the first few points, ii) a major portion of linear relation between two variables, and iii) a 'droop' at the tail end of the distribution indicating the incompleteness of the bibliography (*Brookes, 1969*).

Figure 4. Bradford distribution of journal data

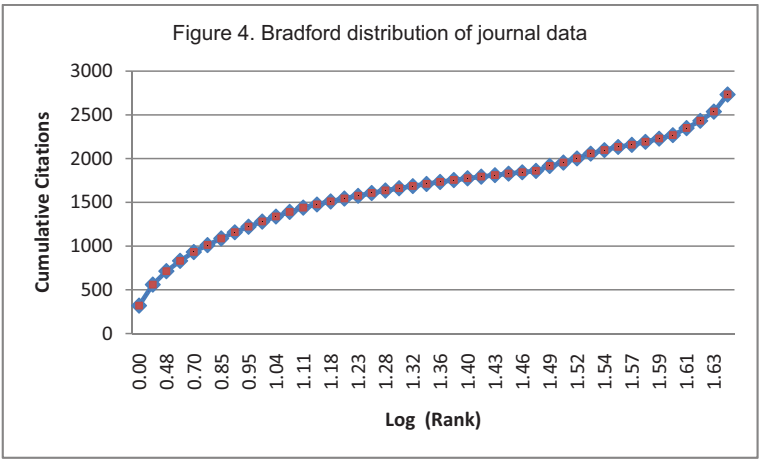


Fig.4 presents the Bradford's distribution of journal citations, in which cumulative citations against the logarithm of journal ranks are plotted. It has been observed that the initial points are not rapidly rise, i.e. the core journals are not sharply identified probably due to inter-disciplinary nature of research

covered in the Journal. However, citation frequency of first 5 journals is reasonably high and may be considered as central core. So it appears that the nature of this curve is far beyond the characteristics of classical S plot.

6.7 Identifying Core Journals

In any domain of research, a few journals are frequently referred by the researchers to conceptualize their new ideas, thereby establish a semantic relation between the subject and journals. These highly cited journals usually makes their influence in a particular domain of research seem to be *core journals*, otherwise may be considered as central set of journals. The core journals always contain higher concentration of relevant articles on particular areas of research in a discipline, thus helps library managers for effective collection development and judicial resource allocation. In fact, identification of core journals in statistical science could enable the researchers, library managers, academic administrators, faculty members of the leading schools of the subject concerned to formulate necessary collection development and management policies in a realistic way.

Table 7 presents the highly cited journals of statistics (along with Impact Factor and country of origin) appended in *Sankhya* during the study period. Table shows that top 12 journal titles have contributed more than 50% of the total citations; thus may be treated as preferred channels of communication in statistical science research. Among the journals, top 5 titles belongs to the most productive zone (read as nuclear zone) might be regarded as core journals and occurs maximum density referred by *Sankhya*. These are *Annals of Statistics* (319), *Journal of the American Statistical Association* (240), *Biometrika* (154), *Journal of the Royal Statistical Society - B* (118), *Annals of Mathematical Statistics* (101). However, a few other journals those exerts adequate influence in the field are *Statistics and Probability Letters* (79), *Journal of Statistical Planning and Inference* (76), *Journal of Multivariate Analysis* (71), *Annals of Probability* (63), and *Technometrics* (59) respectively.

The study also reveals that source journal itself has become a part of the highly cited list of titles. Almost all the titles belong to the coverage of the Journal Citation Report (JCR), thus reflecting the pedigree and prestige of *Sankhya* as renowned scholarly journal. Therefore, the source journals referred highly cited journals; often with high impact factor and originated from different countries, thus endorsed a sort of internationality of the *Sankhya*.

Table 7. Highly cited journals of Statistics (top twenty)

Rank	Journal name (Abbreviation)	ISSN	Freq. of Citation	Impact Factor$^{\Psi}$	Country
1	The Annals of Statistics (AoS)	0090-5364	319	2.307	USA
2	Journal of the American Statistical Association (JASA)	0162-1459	240	2.394	USA
3	Biometrika (Bka)	0006-3444	154	1.405	UK
4	Journal of the Royal Statistical Society. Series B (JRSB)	1369-7412	118	2.835	UK
5	Annals of Mathematical Statistics (AoMS)	0003-4851	101	NC	USA
6	Statistics and Probability Letters (SPL)	0167-7152	79	0.445	Netld.
7	Journal of Statistical Planning and Inference (JSPI)	0378-3758	76	0.679	Netld
8	Journal of Multivariate Analysis (JMA)	0047-259X	71	0.738	USA

9	Annals of Probability (AoP)	0091-1798	63	1.278	USA
10	Technometrics (Tech)	0040-1706	59	1.071	USA
11	Biometrics (Bcs)	0006-341X	58	1.970	UK
12	Sankhya: The Indian Journal of Statistics – A (SnkA)	0581-572X	52	NC	India
13	Communications in Statistics: Theory & Methods (CSTM)	0361-0926	49	0.324	USA
14	Journal of Applied Probability (JAP)	0021-9002	37	0.739	UK
15	Annals of the Institute of Statistical Mathematics (AISM)	0020-3157	34	0.565	Netld.
16	Canadian Journal of Statistics (CJS)	0319-5724	33	0.589	Canada
17	Econometrica (Ecnt)	0012-9682	32	3.865	UK
18	Statistica Sinica (Ssin)	1017-0405	31	0.699	Taiwan
19	Scandinavian Journal of Statistics (SJS)	0303-6898	29	1.268	UK
20	Statistical Science (Ssci)	0883-7237	26	2.135	USA

$^{\Psi}$ *Impact factors have drawn from JCR-2010.* NC = *The data not covered in JCR.*

Worthy to note that, almost all the volumes of these highly cited journals enumerated here (Table-7) are available in the Institute library and still continuing within the current subscription list (both print and online), thereby justifies the collection development policy of the Indian Statistical Institute Library, Kolkata.

7. CONCLUSION

Citation analysis, despite certain ambiguities has often been considered as convenient and reliable measure for research assessment; thus citation behavior of this scholarly journal could be useful to characterize the statistical science research in various lights. Certainly this study informs about the research orientation in the field, thereby signifies the influence of this channel of communication. The analyses presented in this study have permitted many conclusions of broad generality on statistical science research and in particular to *Sankhya*. However, findings obtained in the present study reinforce the findings of the author's earlier study (*Das & Pal, 2012*) and creates a comprehensive portrait of the source journal.

Here the study reveals that a total of 3750 cited references appended in 199 citing articles published in Sankhya during 2003 to 2007. It appears that every issue published approximately 10 articles and each article has an average of 18.84 cited references. In fact the numbers of articles and corresponding citations have decreased gradually over the years, except in 2007. Thus, average citation per article is fairly impressive, indicates that researchers followed a rigorous review of literatures to report their research findings.

However, the contributors have been found to use a wide variety of source materials for reporting their research papers; comprising of journal articles, monographs, conference papers, preprints, technical papers, thesis, dissertations, and also user-guides, audit reports, annual reports, became into existent. Journal articles get cited predominantly than monographs and the use of web-resources is likely very poor cited. Thus citation behavior of Sankhya exhibits a close resemblance with the usual practice in hard science journals. Again, the Journal has found an average of five keywords to represent the thought-contents in each article.

It has been found that a total of 372 unique journal-titles received a total of 2732 citations, thereby journal citation density has been derived as 7.35, is quite significant. Testing of Bradford's law identifies a core list of journals, though the number of journals in the third zone far outnumbers the expected value (125), thus does not fit into Bradford's distribution, and the nature of the curve is far beyond the characteristics of classical S plot. In fact top 12 journal titles have contributed more than 50% of the total citations, which are consistently get cited in *Sankhya* viz. *Annals of Statistics*, *Journal of the American Statistical Association*, *Biometrika*, *Journal of the Royal Statistical Society (B)*, *Annals of Mathematical Statistics*, *Statistics and Probability Letters*, *Journal of Statistical Planning and Inference*, *Journal of Multivariate Analysis*, *Annals of Probability*, *Technometrics*, *Biometrics, Sankhya (A),* etc.

However, the citation frequency of top 5 journals (central core) is reasonably high and belongs to the coverage of the Journal Citation Report (JCR), thus reflecting the pedigree and prestige of *Sankhya* as renowned scholarly journal. It has been found that *Sankhya* usually appended highly reputed journals (often with high impact factor) originated from various countries; thereby endorsed a flavour of internationality. Therefore, Sankhya could stake claim as one of the most authoritative channel of communication for scholarly literatures on statistics and allied areas of research.

All these highly cited titles enumerated in the study expected to be available in a special library dealing with Statistical Science research, in order to maintain a sustainable collection development. Notably, the Institute library (ISI) holds all such back volumes and still continuing in its' current subscription list; hence justifies the collection development policy of the Indian Statistical Institute Library, Kolkata.

In no doubt this study has revealed much information in multiple dimensions, which might be useful to the scholars, library managers, and policy makers to formulate strategies and provide adequate support towards the growth and development of statistical research centers. The authors firmly believe that, as a whole the present study would be helpful in judicious resources allocation, making policies for collection development in library and information centers, as well as could be useful to statistical researchers in choosing their desired channel of further research communications.

8. REFERENCES

Bornmann L., Daniel H. D., 2008, What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*. **64** (1), 45-80.

Bradford, S. C., 1985, Sources of information on specific subjects. *Journal of Information Science.* **10**, 176-180.

Brookes, B.C., 1969, Bradford's law and bibliography of science. *Nature.* **232** (5223), 955.

Das, P. K., Pal, J. K., 2012, Scientometric evaluation of Sankhyá – the Indian Journal of Statistics. *Malaysian Journal of Library and Information Science*. **17** (2), 83-100.

de Solla Price, D., 1963, *Little science big science*. Columbia University Press, New York.

de Solla Price, D., 1970, *Citation measures of hard science, soft science, technology, and nonscience*, in Communication among Scientists and Engineers, *edited by* C. E. Nelson & D. K. Pollack. Heath, Lexington, MA, 3-22p.

de Solla Price, D., 1986, *Little science big science and beyond*. Columbia University Press, New York, 284p.

Editorial, 1933, *Sankhya.* **1** (1).

Egghe, L., Rousseau, R., 1990, *Introduction to informatics: quantitative methods in library, documentation and information science*. Elsevier Sc., Amsterdam.

Garfield, E., 1972, Citation analysis as a tool in journal evaluation. *Science,* **178** (4060), 471-479.

Garfield, E., 2007, The evolution of the Science Citation Index. *International Microbiology*. **10**, 65-69.

Garfield, E., et al. 1978, *Citatation data as science indicators,* in Towards a metric of science, edited by Y. Elkana. John Wiley, New York, 179-207p.

Gross, P.L.K., Gross, E.M., 1927, College libraries and chemical education. *Science*. **66**, 385-389.

Johnson, B., 2000, Environment impact: a preliminary citation analysis of local faculty in a new academic program in environmental and human health applied to collection development in an academic library. *Library Philosophy and Practice*. **2** (2), 1-24.

MathSciNet – American Mathematical Society (AMS), Providence, USA. Available at http://www.ams.org/mathscinet (Accessed on 24 February 2011).

Rousseau, R., 2008, *Publication and citation analysis as a tool for information retrieval*, in Social Information Retrieval Systems: emerging technologies and applications for searching the web effectively (Chapter-XIII), edited by D. Goh & S. Foo. Information Science Reference (IGI Global), Hershey (PA), 252-267p.

Rousseau, R., Hu, X., 2010, An outgrow index. *Annals of Library and Information Studies*, **57** (3), 287-290.

Rudra, A., 1996, *Prasanta Chandra Mahalanobis: a biography*. Oxford University Press, Delhi, 268-270p.

Sandison, A, 1989, Documentation note: thinking about citation analysis. *Journal of Documentation*. **45** (1), 59-64.

Sankhya website. available at http://sankhya.isical.ac.in (accessed on 12 March 2011)

Small, H.G., 2004, On the shoulders of Robert Merton: towards a normative theory of citation. *Scientometrics*. **60** (1), 71-79.

Stigler, S.S., 1994, Citation patterns in the journals of statistics and probability. *Statistical Science.* **9** (1), 94-108.

White, E.C., 1985, Bibliometrics: from curiosity to convention. *Special Libraries*. **76**, 35-42.

Williams, V.K., Fletcher, C.L., 2006, Materials used by master's students in engineering and implications for collection development: a citation analysis. *Issues in Science and Technology Librarianship* [Online]. **45**. Available at *http://www.istl.org/06-winter/refereed1.html* (Accessed on 22 February 2011).

Zainab, A.N., Anyi, K.W.U., Anuar, N.B., 2009, A single journal study: Malayasian Journal of Computer Science. *Malayasian Journal of Library and Information Science*. **22** (1), 1-18.

Ziman, J. M., 1968, *Public knowledge: an essay concerning social dimension of science*. Cambridge University Press, Cambridge.

~~~~~~~~~~~~~~~~
~~~~~~~~~~~~~~~~